

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 16:35:35 ; Search time 9440.04 Seconds
(without alignments)
10279.455 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catgtgtgcccacacagcaca.....gcaataacattcctaataatc 2052

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 4526729 seqs, 23644849745 residues
Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_scs.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|------------|--------------------|
| 1 | 2052 | 100.0 | 2052 | 6 E40087 | E40087 Plant promo |
| 2 | 2048.8 | 99.8 | 2052 | 6 E40093 | E40093 Plant promo |
| 3 | 2038 | 99.3 | 2056 | 6 E40091 | E40091 Plant promo |
| 4 | 2034 | 99.1 | 2048 | 6 E40090 | E40090 Plant promo |
| 5 | 2032.4 | 99.0 | 2048 | 6 E40089 | E40089 Plant promo |
| 6 | 241.2 | 11.8 | 2042 | 6 AR076817 | AR076817 Sequence |
| 7 | 241.2 | 11.8 | 2042 | 6 E15125 | E15125 Promoter. 7 |
| 8 | 196.6 | 9.6 | 247 | 6 AR076816 | AR076816 Sequence |
| 9 | 196.6 | 9.6 | 247 | 6 E15124 | E15124 Promoter. 7 |
| 10 | 195.6 | 9.5 | 246 | 6 AR146852 | AR146852 Sequence |
| 11 | 195.6 | 9.5 | 246 | 6 E55065 | E55065 Plant promo |
| 12 | 174 | 8.5 | 8005 | 8 DCA18706 | Y18706 Daucus caro |
| 13 | 162.4 | 7.9 | 2831 | 6 BD188672 | BD188672 Promotor |
| 14 | 162.4 | 7.9 | 2865 | 6 BD188674 | BD188674 Promotor |
| 15 | 162.4 | 7.9 | 4886 | 8 DARGCHS2 | D16255 Carrot gCHS |
| 16 | 119 | 5.8 | 140 | 6 AR146856 | AR146856 Sequence |
| 17 | 119 | 5.8 | 140 | 6 AR146857 | AR146857 Sequence |
| 18 | 119 | 5.8 | 140 | 6 E55069 | E55069 Plant promo |
| 19 | 119 | 5.8 | 140 | 6 E55070 | E55070 Plant promo |

| | | | | | |
|------|-------|-----|--------|-------------|--------------------|
| c 20 | 106.6 | 5.2 | 250029 | 3 AE014839 | AE014839 Plasmidiu |
| c 21 | 103.6 | 5.0 | 172816 | 9 AC093899 | AC093899 Homo sapi |
| c 22 | 101.8 | 5.0 | 258658 | 3 AE014832 | AE014832 Plasmidiu |
| c 23 | 100.4 | 4.9 | 348174 | 3 CR382399 | CR382399 Plasmidiu |
| c 24 | 99.4 | 4.8 | 8056 | 6 AX599046 | AX599046 Sequence |
| c 25 | 99.2 | 4.8 | 205130 | 2 AC105425 | AC105425 Homo sapi |
| c 26 | 97.6 | 4.8 | 176666 | 2 CR392006 | CR392006 Dario rer |
| c 27 | 97 | 4.7 | 14867 | 3 AE001398 | AE001398 Plasmidiu |
| c 28 | 96.6 | 4.7 | 67970 | 3 PFMAL1P3 | AL031746 Plasmidiu |
| c 29 | 96.4 | 4.7 | 4601 | 3 DMU11584 | U11584 Drosophila |
| c 30 | 96.4 | 4.7 | 19517 | 3 DMU37541 | U37541 Drosophila |
| c 31 | 95.8 | 4.7 | 64394 | 9 AL928596 | AL928596 Human DNA |
| c 32 | 95.2 | 4.6 | 313050 | 3 PFA929352 | AL929352 Plasmidiu |
| c 33 | 93.8 | 4.6 | 143331 | 9 AC091214 | AC091214 Homo sapi |
| c 34 | 93 | 4.5 | 349980 | 6 AX344555 | AX344555 Sequence |
| c 35 | 92.6 | 4.5 | 180629 | 2 CR388025 | CR388025 Dario rer |
| c 36 | 92.6 | 4.5 | 349751 | 3 PFMAL4P3 | AL035476 Plasmidiu |
| c 37 | 92.4 | 4.5 | 261771 | 2 CR356223 | CR356223 Dario rer |
| c 38 | 91.8 | 4.5 | 8056 | 6 AX598900 | AX598900 Sequence |
| c 39 | 91.6 | 4.5 | 8056 | 6 AX599046 | AX599046 Sequence |
| c 40 | 91 | 4.4 | 21153 | 8 YSCMTG13 | L3697 Saccharomyc |
| c 41 | 91 | 4.4 | 118642 | 9 AC126283 | AC126283 Homo sapi |
| c 42 | 90.8 | 4.4 | 86826 | 3 PFMAL3P5 | AL034556 Plasmidiu |
| c 43 | 90.6 | 4.4 | 155106 | 9 AC104069 | AC104069 Homo sapi |
| c 44 | 90.6 | 4.4 | 178670 | 9 AC104073 | AC104073 Homo sapi |
| c 45 | 90.2 | 4.4 | 105238 | 9 AC011458 | AC011458 Homo sapi |

ALIGNMENTS

RESULT 1
E40087
LOCUS E40087 2052 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40087
VERSION E40087.1 GI:18627203
KEYWORDS JP 2000166577-A/1.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2052)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 1 20-JUN-2000;
COMMENT SUMITOMO CHEM CO LTD
OS Daucus carota L.
PN JP 2000166577-A/1
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR
PI SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC
C12N15/00,
PC C12N5/00,(C12N5/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter (1)..(2052).
FEATURES
source Location/Qualifiers
1..2052
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 100.0%; Score 2052; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 3.3e-310;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGTGTGCCCCACACATAGGCGCTTTGGTTGAGAGAGCAGAGCTGCTTCGA 60
DB 1 CATGTGTGCCCCACACATAGGCGCTTTGGTTGAGAGAGCAGAGCTGCTTCGA 60

QY 61 CTTCTTCTCTTTGACCTGTTGTATAAAGAGTAGAATAATATTTTAAAGCTGCGAA 120
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 CTTCTTCTCTTTGACCTGTTGTATAAAGAGTAGAATAATATTTTAAAGCTGCGAA 120
QY 121 TACTAACTTCTCTCACAACCTTCGGCTTCTTTTCCAAACACCTTTATTAACCTTTTACT 180
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 TACTAACTTCTCTCACAACCTTCGGCTTCTTTTCCAAACACCTTTATTAACCTTTTACT 180
QY 181 TCTCAATTTCTACTCCACCTTCTTTGCTATAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 TCTCAATTTCTACTCCACCTTCTTTGCTATAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240
QY 241 AACGCCCTCAATAAAGAGTCAATTCATAAATGTATCTTTCAATTTTAGGATTAACAATACGT 300
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 AACGCCCTCAATAAAGAGTCAATTCATAAATGTATCTTTCAATTTTAGGATTAACAATACGT 300
QY 301 GAAACAGGGTATTTTAAACGTGTCAACAAATTTCTAAATAATTTTACCTGGCGGTGAACA 360
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 GAAACAGGGTATTTTAAACGTGTCAACAAATTTCTAAATAATTTTACCTGGCGGTGAACA 360
QY 361 CCGTCTTCCAGATTAATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 CCGTCTTCCAGATTAATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTCGCATGC 420
QY 421 AGGACGACTTAGGTGAATACACATTTGTACTGTAGCTCTTTAAACAAAGCAAGTGGTTC 480
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 AGGACGACTTAGGTGAATACACATTTGTACTGTAGCTCTTTAAACAAAGCAAGTGGTTC 480
QY 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACCAACTCTATCCAGTACTATACCTTT 540
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACCAACTCTATCCAGTACTATACCTTT 540
QY 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATAATGCCCCATCCCAAGGATAAGT 600
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATAATGCCCCATCCCAAGGATAAGT 600
QY 601 AAAATTCGGTTTAAACAGTTGTTTAAATATATATCTTTTACCTTACAGAGGATATTCGT 660
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 AAAATTCGGTTTAAACAGTTGTTTAAATATATATCTTTTACCTTACAGAGGATATTCGT 660
QY 661 AATATCTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 AATATCTTTTAGACGACAAGAGACTTAGGTCAAAAATGGACGCTGGTAAACAGCCTAGACT 720
QY 721 TGGTCACTGATAAATAGATAAATGTTAGTATTAATATAGTAGATCTCAATGACATTAATA 780
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 TGGTCACTGATAAATAGATAAATGTTAGTATTAATATAGTAGATCTCAATGACATTAATA 780
QY 781 ATTAGAGCTATTAATTTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 ATTAGAGCTATTAATTTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTA 840
QY 841 AAAACAAGAGCTTGTCTGTGTGTTTGTAGTTGTGTGAGCTCATTTCTTTTAAAGTAATG 900
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
841 AAAACAAGAGCTTGTCTGTGTGTTTGTAGTTGTGTGAGCTCATTTCTTTTAAAGTAATG 900
QY 901 TAACTGATCTAAAGCACAATGAATAATTTAGTACAGGTTAAACCTTTTACAAGAAATTTATA 960
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
901 TAACTGATCTAAAGCACAATGAATAATTTAGTACAGGTTAAACCTTTTACAAGAAATTTATA 960
QY 961 TTAACGAAATCAATTTTATACATGCTCTCGGCTGTCATTTAATATAGGATCACTTAC 1020
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
961 TTAACGAAATCAATTTTATACATGCTCTCGGCTGTCATTTAATATAGGATCACTTAC 1020
QY 1021 TGATCATCCATTTAAACCTTGTAAACAAATTTCAATGAGATAAATAATCTTTACAATGAA 1080
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1021 TGATCATCCATTTAAACCTTGTAAACAAATTTCAATGAGATAAATAATCTTTACAATGAA 1080
QY 1081 AAGAGGACAATGTCTCTTTGAAACAAACAAATAGGTACTCCCTCGCTCCCTCTGAAATGT 1140
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1081 AAGAGGACAATGTCTCTTTGAAACAAACAAATAGGTACTCCCTCGCTCCCTCTGAAATGT 1140
QY 1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200

DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1141 ATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
QY 1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGACCCACCACCAATATATAAATGTAGATTTAG 1260
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1201 AAAGAGAAAGAAAAGTGGGTAAAGTAGCGGACCCACCACCAATATATAAATGTAGATTTAG 1260
QY 1261 AAAAGTAGTTGAAAAGTAGTGGGTGGGTGGGATTTTATATTTAAAAATTTACTATTTTG 1320
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1261 AAAAGTAGTTGAAAAGTAGTGGGTGGGTGGGATTTTATATTTAAAAATTTACTATTTTG 1320
QY 1321 AGAAAGTTTTCAGAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAA 1380
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1321 AGAAAGTTTTCAGAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAA 1380
QY 1381 TAAATGGGACAGAGGAGTAAATACCTTTATCATATATAAAATTTTGTATTTTGATTTTCA 1440
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1381 TAAATGGGACAGAGGAGTAAATACCTTTATCATATATAAAATTTTGTATTTTGATTTTCA 1440
QY 1441 TAAGATTATAAATCTATGTTTATAATGATAATATAATTTTAAAAATAAATACTATATAAT 1500
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1441 TAAGATTATAAATCTATGTTTATAATGATAATATAATTTTAAAAATAAATACTATATAAT 1500
QY 1501 CTGATTAGTCGATTACCGCCTTTTATATTTTCAATACAGTAAATGATAAATAAATCAG 1560
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1501 CTGATTAGTCGATTACCGCCTTTTATATTTTCAATACAGTAAATGATAAATAAATCAG 1560
QY 1561 TTATCTGAAAAGCAAAATAATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAAGTTTCATG 1620
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1561 TTATCTGAAAAGCAAAATAATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAAGTTTCATG 1620
QY 1621 TGTATTTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTTTCAGA 1680
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1621 TGTATTTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTTTCAGA 1680
QY 1681 AATTTAAAAATAAATTTATTGAGCATGGGAAGTTTCAAGGCATCATTTGAGCAGCAGTACT 1740
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1681 AATTTAAAAATAAATTTATTGAGCATGGGAAGTTTCAAGGCATCATTTGAGCAGCAGTACT 1740
QY 1741 GTTTGAAACAATGTATGTCGGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATAT 1800
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1741 GTTTGAAACAATGTATGTCGGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATAT 1800
QY 1801 GCATTTCTAGATAATCATCTTTTCAAAATTTTCAACAAACAGAGCTTTAACTTTCTTCAACG 1860
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1801 GCATTTCTAGATAATCATCTTTTCAAAATTTTCAACAAACAGAGCTTTAACTTTCTTCAACG 1860
QY 1861 GATTGGAATCCCTTTTCTAAACTTTTAAAAATAAATAAATGATTTATTGTAATATTATC 1920
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1861 GATTGGAATCCCTTTTCTAAACTTTTAAAAATAAATAAATGATTTATTGTAATATTATC 1920
QY 1921 AACACCTCAACATTTGATTTAGCGTACTATATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1921 AACACCTCAACATTTGATTTAGCGTACTATATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
QY 1981 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTAATTTCTCAGCAATTAAC 2040
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1981 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTACTTAATTTCTCAGCAATTAAC 2040
QY 2041 ATTTCTAAATATC 2052
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2041 ATTTCTAAATATC 2052

RESULT 2

E40093
LOCUS E40093 2052 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40093
VERSION E40093.1 GI:18627209
KEYWORDS JP 2000166577-A/7.
SOURCE unidentified
ORGANISM unidentified

| | |
|---|---|
| unclassified. | |
| REFERENCE | 1 (bases 1 to 2052) |
| AUTHORS | Nishikawa,S. and Oeda,K. |
| TITLE | Plant promoter and terminator |
| JOURNAL | Patent: JP 200016577-A 7 20-JUN-2000; |
| COMMENT | SUMITOMO CHEM CO LTD |
| OS | Daucus carota L. |
| PN | JP 200016577-A/7 |
| PD | 20-JUN-2000 |
| PF | 01-OCT-1999 JP 1999281475 |
| PR | |
| PI | SATOMI NISHIKAWA, KENJI OEDA |
| PC | C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC |
| CC | C12N5/00, (C12N5/00,C12R1:91) |
| Key | Location/Qualifiers |
| FT | promoter (1)..(2052). |
| FEATURES | Location/Qualifiers |
| source | 1..2052 |
| | /organism="unidentified" |
| | /mol_type="genomic DNA" |
| | /db_xref="taxon:32644" |
| ORIGIN | |
| Query Match 99.8%; Score 2048.8; DB 6; Length 2052; | |
| Best Local Similarity 99.9%; Pred. No. 1e-309; 2; Indels 0; Gaps 0; | |
| Matches 2050; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | |
| QY | 1 CATGTGCGCTACAGCAGCATAGGGCCGTGTTGGTTGAGAGAGCAGAGCTGCTCTGA 60 |
| DB | 1 CATGTGCGCTACAGCAGCATAGGGCCGTGTTGGTTGAGAGAGCAGAGCTGCTCTGA 60 |
| QY | 61 CTTCTCTCTTTGACCTGTTGTATAAGAGTAGAATATTTTAAAGAGCTCGAA 120 |
| DB | 61 CTTCTCTCTTTGACCTGTTGTATAAGAGTAGAATATTTTAAAGAGCTCGAA 120 |
| QY | 121 TACTAACTTCTCTCACAACTTCCGCTTCTTTCCAAACACTTTATTAACTTTTACT 180 |
| DB | 121 TACTAACTTCTCTCACAACTTCCGCTTCTTTCCAAACACTTTATTAACTTTTACT 180 |
| QY | 181 TCTCATTTCTACCTGCTTGTGCTATAAGCAAGAAATCACTCTTTTAAAGCTAACCCA 240 |
| DB | 181 TCTCATTTCTACCTGCTTGTGCTATAAGCAAGAAATCACTCTTTTAAAGCTAACCCA 240 |
| QY | 241 AAGGCGCTCAATAAAGATCATTATATAAGTAGTATCTTTCAATTTTAGGATACATACGT 300 |
| DB | 241 AAGGCGCTCAATAAAGATCATTATATAAGTAGTATCTTTCAATTTTAGGATACATACGT 300 |
| QY | 301 GAACAGGGTTATTTTAAAGTGTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA 360 |
| DB | 301 GAACAGGGTTATTTTAAAGTGTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA 360 |
| QY | 361 CGGCTTCCAAGATAATATTTTAAATTTGTAGCCTCCCTTTTAAACCAAAATTCGATGC 420 |
| DB | 361 CGGCTTCCAAGATAATATTTTAAATTTGTAGCCTCCCTTTTAAACCAAAATTCGATGC 420 |
| QY | 421 AGGACCACTTAGGTGATACATTTGCTGCTGAGTCTTTTAAACCAAGCAAGTGGTTC 480 |
| DB | 421 AGGACCACTTAGGTGATACATTTGCTGCTGAGTCTTTTAAACCAAGCAAGTGGTTC 480 |
| QY | 481 ATGCTCAGGCATCAAAATTTGACAAAACCGCAGACACACTCTATCCAGCTACTATCTTT 540 |
| DB | 481 ATGCTCAGGCATCAAAATTTGACAAAACCGCAGACACACTCTATCCAGCTACTATCTTT 540 |
| QY | 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAGGATAAGT 600 |
| DB | 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAGGATAAGT 600 |
| QY | 601 AAAATTCGGTTTAAACAGTTGTTAATATATATGTTTACACTTACAAGAGGATATCGT 660 |
| DB | 601 AAAATTCGGTTTAAACAGTTGTTAATATATATGTTTACACTTACAAGAGGATATCGT 660 |
| QY | 661 AATACCTTTTAGACGACAGAGACTTTAGGTCAAAAAATGGACGCTGGTAAACAGCCTAGACT 720 |
| DB | 661 AATACCTTTTAGACGACAGAGACTTTAGGTCAAAAAATGGACGCTGGTAAACAGCCTAGACT 720 |
| QY | 721 TGGTCACTGATAAATAGATAAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTAA 780 |
| DB | 721 TGGTCACTGATAAATAGATAAATTTGTTAGTATAATATAGTAGGATCTACAATGACATTAA 780 |
| QY | 781 ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAAAGCAGGTA 840 |
| DB | 781 ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAAAGCAGGTA 840 |
| QY | 841 AAAACAAGAGCTTGTCTGCTGTGTTTGTAGTTGTGAGCTCATTTCTTTAAAGTAATG 900 |
| DB | 841 AAAACAAGAGCTTGTCTGCTGTGTTTGTAGTTGTGAGCTCATTTCTTTAAAGTAATG 900 |
| QY | 901 TAAACTGATCTAAAGCAGACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA 960 |
| DB | 901 TAAACTGATCTAAAGCAGACATAGAAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTTATA 960 |
| QY | 961 TTTAAACGAAAATCAATTTTATAACATGCTCTCGGCTGTCAATTAATAGGGATCACTTAC 1020 |
| DB | 961 TTTAAACGAAAATCAATTTTATAACATGCTCTCGGCTGTCAATTAATAGGGATCACTTAC 1020 |
| QY | 1021 TGATCATCCATTAAAAACCTTGTAAAAACAAATTTCAATGAGATAAAATATCTTACAATGAA 1080 |
| DB | 1021 TGATCATCCATTAAAAACCTTGTAAAAACAAATTTCAATGAGATAAAATATCTTACAATGAA 1080 |
| QY | 1081 AAGAAGGACAATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCCTCTGAAATGT 1140 |
| DB | 1081 AAGAAGGACAATGTCTCTTTGAAAAAACAATAGGTACTCCCTCCGCTCCTCTGAAATGT 1140 |
| QY | 1141 ATACATATGGAATTTGAGACCGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200 |
| DB | 1141 ATACATATGGAATTTGAGACCGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200 |
| QY | 1201 AAGAAGAAAGAAAGTGGTAAAGTAGCGGGACCCACCAATATATAATTTGATAGATTTAG 1260 |
| DB | 1201 AAGAAGAAAGAAAGTGGTAAAGTAGCGGGACCCACCAATATATAATTTGATAGATTTAG 1260 |
| QY | 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATATAAATAATTTACTATTTTG 1320 |
| DB | 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAATAATTTACTATTTTG 1320 |
| QY | 1321 AGAAAGTTTGTAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380 |
| DB | 1321 AGAAAGTTTGTAAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380 |
| QY | 1381 TAAATGGGACAGAGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGAATTTCA 1440 |
| DB | 1381 TAAATGGGACAGAGGAGTAAATACCTTTATGATATATAAATTTTGTATTTTGAATTTCA 1440 |
| QY | 1441 TAAAGATTATAAATCTATGTTTATAATGATAATAATATTTTAAATAATATCTATATAAT 1500 |
| DB | 1441 TAAAGATTATAAATCTATGTTTATAATGATAATAATATTTTAAATAATATCTATATAAT 1500 |
| QY | 1501 CTGATTAGTCGATTACCGCTTTTATATTTTACATACTAGTATATGAATAAATCAG 1560 |
| DB | 1501 CTGATTAGTCGATTACCGCTTTTATATTTTACATACTAGTATATGAATAAATCAG 1560 |
| QY | 1561 TTATCTGAAAACCAATAATATCTTTGTAAAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1620 |
| DB | 1561 TTATCTGAAAACCAATAATATCTTTGTAAAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1620 |
| QY | 1621 TGTATTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTTATTTTGTTCAGA 1680 |
| DB | 1621 TGTATTCAATAGTTTTTAATATAAAGTAAATTTTAAATTAATTTGTTTATTTTGTTCAGA 1680 |
| QY | 1681 AATTTTAAATAAATTTTATGAGCATGGAGTTTCAGGGCATCATTTGAGCAGCAGCTAGACT 1740 |
| DB | 1681 AATTTTAAATAAATTTTATGAGCATGGAGTTTCAGGGCATCATTTGAGCAGCAGCTAGACT 1740 |
| QY | 1741 GTTTGAACAATATGATGTCGGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800 |

Db 1741 GTTTGAACAATGATGTCGGGTGACATCATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Qy 1801 GCATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
Db 1801 GCATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTTCAACG 1860
Qy 1861 GATTGGAATCCTTTTCTAAACCTTTTAAATAAAAAAATGCAATATTGTTGAATATTATC 1920
Db 1861 GATTGGAATCCTTTTCTAAACCTTTTAAATAAAAAAATGCAATATTGTTGAATATTATC 1920
Qy 1921 AACACCTCAACATGATGTTAGCTACTATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
Db 1921 AACACCTCAACATGATGTTAGCTACTATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
Qy 1981 ACATCAATCTTTACACCAACCTTGAGCTTTAAATTTTTCTACTATTATCTCAGCAATAAC 2040
Db 1981 ACATCAATCTTTACACCAACCTTGAGCTTTAAATTTTTCTACTATTATCTCAGCAATCAC 2040
Qy 2041 ATTCTAAATATC 2052
Db 2041 ATTCTAAAGATC 2052

RESULT 3

E40091
LOCUS E40091 2056 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40091
VERSION E40091.1 GI:18627207
KEYWORDS JP 2000166577-A/5.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2056)
AUTHORS Nishikawa S. and Oeda K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 5 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.
PN JP 2000166577-A/5
PD 20-JUN-2000
PP 01-OCT-1999 JP 1999281475
PR SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC
C12N15/00,
PC C12N5/00, (C12N5/00,C12R1:91)
CC

FEATURES
source
FH Key Location/Qualifiers
FT promoter (1)..(2056).
1..2056
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 99.3%; Score 2038; DB 6; Length 2056;
Best Local Similarity 99.8%; Pred. No. 5e-308;
Matches 2052; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
Qy 1 CATGTGTCCTTACAGCATAGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGA 60
Db 1 CATGTGTCCTTACAGCATAGGCGCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGA 60
Qy 61 CTTCCTCTCTTTTTCACCTGTTGTATAGAGAGTAGAAATATTTTAAAGCTGCGAA 120
Db 61 CTTCCTCTCTTTTTCACCTGTTGTATAGAGAGTAGAAATATTTTAAAGCTGCGAA 120
Qy 121 TACTAACTTCTCTCACAACCTTCGGCTCTTTTCCAAACACTTTATTAACTTTTACT 180
Db 121 TACTAACTTCTCTCACAACCTTCGGCTCTTTTCCAAACACTTTATTAACTTTTACT 180

Qy 181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAAACCCA 240
Db 181 TCTCATTTCTACTCCACTTCTTTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAAACCCA 240
Qy 241 AACGGCTCAATAAAGATCAATTCATAAATGTAATCTTTCAATTTTAGATTAACATAGCT 300
Db 241 AACGGCTCAATAAAGATCAATTCATAAATGTAATCTTTCAATTTTAGATTAACATAGCT 300
Qy 301 GAACAGGGTATTATTTTAAACGCTGCAACAATTTCTAATAATTTTACCTGGCGGGTGAACA 360
Db 301 GAACAGGGTATTATTTTAAACGCTGCAACAATTTCTAATAATTTTACCTGGCGGGTGAACA 360
Qy 361 CCGTCTTCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTTGCGATGC 420
Db 361 CCGTCTTCAAGATAATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTTGCGATGC 420
Qy 421 AGGACGACTTAGGTGAATACACATTTGTACTGTAGTCTTTTAAACAAAGCAAGTGGTTC 480
Db 421 AGGACGACTTAGGTGAATACACATTTGTACTGTAGTCTTTTAAACAAAGCAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATACTTT 540
Db 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATACTTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGT 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAATAATATGCCCATCCAGGATAAGT 600
Qy 601 AAAATTCGGTTAAACCAAGTGTGTTAATATATATGTTTACACTTACAAGAGATATTCGT 660
Db 601 AAAATTCGGTTAAACCAAGTGTGTTAATATATATGTTTACACTTACAAGAGATATTCGT 660
Qy 661 AATACTTTTACACGACCAAGAGACTTAGGTCAAAATGAGCGCTGTAAACAGCCTAGACT 720
Db 661 AATACTTTTACACGACCAAGAGACTTAGGTCAAAATGAGCGCTGTAAACAGCCTAGACT 720
Qy 721 TGGTCACTGATAAATAGATAAATGTTAGTATAATATAGTAGGATCTACAATGACATTTAA 780
Db 721 TGGTCACTGATAAATAGATAAATGTTAGTATAATATAGTAGGATCTACAATGACATTTAA 780
Qy 781 ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAAACGAGTA 840
Db 781 ATTAGAGCTATTAAATTAAGTTACTAATAAATAGAGAGGTTAGTAAACAGAAACGAGTA 840
Qy 841 AAAACAAGAGCTTGTCTGTGTGTTAGTTGTTGAGCTCATTTCTTTAAAGTAATG 900
Db 841 AAAACAAGAGCTTGTCTGTGTGTTAGTTGTTGAGCTCATTTCTTTAAAGTAATG 900
Qy 901 TAAACTGATCTAAAGCACAATAGAAATTTAGTACAGGTTTAAACCTTTTCAAGAATTTATA 960
Db 901 TAAACTGATCTAAAGCACAATAGAAATTTAGTACAGGTTTAAACCTTTTCAAGAATTTATA 960
Qy 961 TTAACGAAATCAATTTTATAACATGCTCTCGGCTGTCATTATAATAGGATCACTTAC 1020
Db 961 TTAACGAAATCAATTTTATAACATGCTCTCGGCTGTCATTATAATAGGATCACTTAC 1020
Qy 1021 TGATCATCCATTAAAAACCTTGTAAAAAATTTCAATGAGATAAATAATCTTTTACAAATGAA 1080
Db 1021 TGATCATCCATTAAAAACCTTGTAAAAAATTTCAATGAGATAAATAATCTTTTACAAATGAA 1080
Qy 1081 AAGAAGGACATGCTCTTTGAAAAAACAATAGTACTCCCTCGCTCCCTCTGAAATGT 1140
Db 1081 AAGAAGGACATGCTCTTTGAAAAAACAATAGTACTCCCTCGCTCCCTCTGAAATGT 1140
Qy 1141 ATACATATGGATGGACACGAGACTAAGAAAAATGTAATAAGTAACTGTAGATGTAAG 1200
Db 1141 ATACATATGGATGGACACGAGACTAAGAAAAATGTAATAAGTAACTGTAGATGTAAG 1200
Qy 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTTAG 1260
Db 1201 AAAGAGAAAGAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTTAG 1260

QY 1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAAAAATTTACTATTTTG 1320
DB |||||||
1261 AAAAGTAGTTGAAAGTAGTGGGTGGGTGGGATTTTATATATAAAAAATTTACTATTTTG 1320
QY 1321 AGAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
DB |||||||
1321 AGAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAT 1380
QY 1381 TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGAATTCA 1440
DB |||||||
1381 TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGAATTCA 1440
QY 1441 TAAGATTATAAATCTATGTTTAAATGATATAATAATTTTAAATAATACTATATTAAT 1500
DB |||||||
1441 TAAGATTATAAATCTATGTTTAAATGATATAATAATTTTAAATAATACTATATTAAT 1500
QY 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATACCTGAGTAATATGAATAATCAG 1560
DB |||||||
1501 CTGATTAGTCGATTACCGCTTTTATAATTTTACAATACCTGAGTAATATGAATAATCAG 1560
QY 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1620
DB |||||||
1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTTCGGTCAAAATGGGAAGTTTCATG 1620
QY 1621 TGTATTCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTTGTATTTTGTTCAGA 1680
DB |||||||
1621 TGTATTCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTTGTATTTTGTTCAGA 1680
QY 1681 AATTTAAATAATTTATGAGCATGGGAAGTTTCACGGGCATCATTTGAGCAGCACTAGACT 1740
DB |||||||
1681 AATTTAAATAATTTATGAGCATGGGAAGTTTCACGGGCATCATTTGAGCAGCACTAGACT 1740
QY 1741 GTTTGAACAATGATGTCCGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
DB |||||||
1741 GTTTGAACAATGATGTCCGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
QY 1801 GCATTCTAGCTAGTACATCTTTTCAAAATTTTCAACAAACACAGCTTTAACTTTTCTTTC 1856
DB |||||||
1801 GCATTCTAGCTAGTACATCTTTTCAAAATTTTCAACAAACACAGCTTTAACTTTTCTTTC 1856
QY 1857 AACGGATTGAATCCCTTTTCAACTTTTAAATATAAATAATGCAATTTGTAATAAT 1916
DB |||||||
1857 AACGGATTGAATCCCTTTTCAACTTTTAAATATAAATAATGCAATTTGTAATAAT 1916
QY 1861 AACGGATTGAATCCCTTTTCAACTTTTAAATATAAATAATGCAATTTGTAATAAT 1920
DB |||||||
1861 AACGGATTGAATCCCTTTTCAACTTTTAAATATAAATAATGCAATTTGTAATAAT 1920
QY 1917 TATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGGTGCTCTTGGTGTCTACTA 1976
DB |||||||
1917 TATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGGTGCTCTTGGTGTCTACTA 1976
QY 1921 TATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGGTGCTCTTGGTGTCTACTA 1980
DB |||||||
1921 TATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGGTGCTCTTGGTGTCTACTA 1980
QY 1977 TCATCATCAATCTTTACACCAAAACCTTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAA 2036
DB |||||||
1977 TCATCATCAATCTTTACACCAAAACCTTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAA 2036
QY 1981 TCATCATCAATCTTTACACCAAAACCTTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAA 2040
DB |||||||
1981 TCATCATCAATCTTTACACCAAAACCTTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAA 2040
QY 2037 TAACATCTTAATATC 2052
DB |||||||
2037 TAACATCTTAATATC 2052
QY 2041 TAACATCTTAATATC 2056
DB |||||||
2041 TAACATCTTAATATC 2056

RESULT 4

E40090
LOCUS 2048 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40090
VERSION E40090.1 GI:18627206
KEYWORDS JP 2000166577-A/4.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2048)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 4 20-JUN-2000;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L.

PN JP 2000166577-A/4
PD 20-JUN-2000
PP 01-OCT-1999 JP 1999281475
PR SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09, A01H5/00, C12N1/21, C12N5/10// (C12N5/10, C12R1:91), PC
C12N15/00,
PC C12N5/00, (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT promoter Location/Qualifiers
1..2048
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
FEATURES
source
ORIGIN
Query Match 99.1%; Score 2034; DB 6; Length 2048;
Best Local Similarity 99.8%; Pred. No. 2.1e-307;
Matches 2048; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
QY 1 CATGTGTCCTACAGCATAGGGCTGTGGTTGAGAGAGAGAGAGCTGCTCTCTGA 60
DB 1 CATGTGTCCTACAGCATAGGGCTGTGGTTGAGAGAGAGAGAGCTGCTCTCTGA 60
QY 61 CTTCTCTCTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAGCTCGAA 120
DB 61 CTTCTCTCTTTGACCTGTTGTATAAAGAGTAGAATAATTTTAAAGCTCGAA 120
QY 121 TACTAACTCTCTCACAACTTCGCTTCCTTTTCCAAAACCTTTTAACTTTTACT 180
DB 121 TACTAACTCTCTCACAACTTCGCTTCCTTTTCCAAAACCTTTTAACTTTTACT 180
QY 181 TCTCATTTCTCTCACTCTTTGCTATAAGCAAGAAATCACTTTTAAAGCTAACCA 240
DB 181 TCTCATTTCTCTCACTCTTTGCTATAAGCAAGAAATCACTTTTAAAGCTAACCA 240
QY 241 AACGGCTCAATAAAGATCATTTCAATAATGTATCTTTCAATTTTAAAGTAACAATACGT 300
DB 241 AACGGCTCAATAAAGATCATTTCAATAATGTATCTTTCAATTTTAAAGTAACAATACGT 300
QY 301 GAACAGGGTATTTTAAAGTGTCAACAAATTTCTAATAATTTTAACTTCGCGGTGAACA 360
DB 301 GAACAGGGTATTTTAAAGTGTCAACAAATTTCTAATAATTTTAACTTCGCGGTGAACA 360
QY 361 CGGTCTTCAAGATAATAATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTTCCGATGC 420
DB 361 CGGTCTTCAAGATAATAATTTTAAATTTTGTAGCTCCCTTTTAAACCAATTTCCGATGC 420
QY 421 AGGAGACTTAGTGTAATACATCTGTAGTCTTTTAAACCAAGCAAGTGGTTC 480
DB 417 AGGAGACTTAGTGTAATACATCTGTAGTCTTTTAAACCAAGCAAGTGGTTC 476
QY 481 ATGCTCAGCCATCAAAATTTGACAAAACCGGACACAACTCTATCCACGCTACTATCTTT 540
DB 477 ATGCTCAGCCATCAAAATTTGACAAAACCGGACACAACTCTATCCACGCTACTATCTTT 536
QY 541 TGGCGAATGCTTCTCAAAATGTTTTTATATGTAATAATTCGCCATCCAGGTAAGT 600
DB 537 TGGCGAATGCTTCTCAAAATGTTTTTATATGTAATAATTCGCCATCCAGGTAAGT 596
QY 601 AAAATTTCCGCTTTAACCCAGTTTGTATATATATATGTTTACACTTACAAGAGGATATTCGT 660
DB 597 AAAATTTCCGCTTTAACCCAGTTTGTATATATATATGTTTACACTTACAAGAGGATATTCGT 656
QY 661 AATACTTTTAGCAGCAAGAGACTTTAGGTCAAAAATGGAGCTGGTGGTAAACAGCCTAGACT 720
DB 657 AATACTTTTAGCAGCAAGAGACTTTAGGTCAAAAATGGAGCTGGTGGTAAACAGCCTAGACT 716
QY 721 TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACATGACATTA 780
DB 717 TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACATGACATTA 776

| | | | |
|----|------|---|------|
| Qy | 781 | ATTAGAGCTATTAAATTAAGTTACTTAATAAATAAGAGAGGTTAGTTAAACAGAAAGCAGGTA | 840 |
| Db | 777 | ATTAGAGCTATTAAATTAAGTTACTTAATAAATAAGAGAGGTTAGTTAAACAGAAAGCAGGTA | 836 |
| Qy | 841 | AAAACAGAGCTTGCTGCTGTGTGTTAGTTGTTGTGAGCTCATTTCTTTAAAGTAATG | 900 |
| Db | 837 | AAAACAGAGCTTGCTGCTGTGTGTTAGTTGTTGTGAGCTCATTTCTTTAAAGTAATG | 896 |
| Qy | 901 | TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAATTTATA | 960 |
| Db | 897 | TAAACTGATCTAAAGCACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAATTTATA | 956 |
| Qy | 961 | TTAAACGAAATCAATTTTATAACATGCTCTCGGCTGTCAATTAATAAGGATCACTTAC | 1020 |
| Db | 957 | TTAAACGAAATCAATTTTATAACATGCTCTCGGCTGTCAATTAATAAGGATCACTTAC | 1016 |
| Qy | 1021 | TGATCATCCATTAAACCTTTTAAACAAATTCAAATGAGATAAAATATCTTACAATGAA | 1080 |
| Db | 1017 | TGATCATCCATTAAACCTTTTAAACAAATTCAAATGAGATAAAATATCTTACAATGAA | 1076 |
| Qy | 1081 | AAGAAGGACAAATGTCTCTTTGAAAAACAAATAGTACTCCCTCCGCTCCCTCTGAAATGT | 1140 |
| Db | 1077 | AAGAAGGACAAATGTCTCTTTGAAAAACAAATAGTACTCCCTCCGCTCCCTCTGAAATGT | 1136 |
| Qy | 1141 | ATACATATGGATTGGACACGGAGCTTAAGAAAAATGTATAAGTAATGTATAGATGTAAGAAG | 1200 |
| Db | 1137 | ATACATATGGATTGGACACGGAGCTTAAGAAAAATGTATAAGTAATGTATAGATGTAAGAAG | 1196 |
| Qy | 1201 | AAACAGAAAAAGAAAGTGGTAAAGTACGGGACCCACCATAATATAATTCATAGATTTAG | 1260 |
| Db | 1197 | AAACAGAAAAAGAAAGTGGTAAAGTACGGGACCCACCATAATATAATTCATAGATTTAG | 1256 |
| Qy | 1261 | AAAAAGTAGTTGAAAGTAGTGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTG | 1320 |
| Db | 1257 | AAAAAGTAGTTGAAAGTAGTGGTGGGTGGGATTTTATATATATAAAAAATTTACTATTTTG | 1316 |
| Qy | 1321 | AGAAAGTTTGGAAATGTATAGAAATGATGGGACATCCATAAAGGAAAGTGATAGAAT | 1380 |
| Db | 1317 | AGAAAGTTTGGAAATGTATAGAAATGATGGGACATCCATAAAGGAAAGTGATAGAAT | 1376 |
| Qy | 1381 | TAAATGGACAGAGGGAGTAATACCTTTATGATATATAAAATTTTTGTATATTTTGATTTC | 1440 |
| Db | 1377 | TAAATGGACAGAGGGAGTAATACCTTTATGATATATAAAATTTTTGTATATTTTGATTTC | 1436 |
| Qy | 1441 | TAAGATTATAAATCTATGTTTATAATGATAATAATTTTAAAAATATACTATATTAAAT | 1500 |
| Db | 1437 | TAAGATTATAAATCTATGTTTATAATGATAATAATTTTAAAAATATACTATATTAAAT | 1496 |
| Qy | 1501 | CTGATTAGTCGATTTACCGCTTTTATAATTTTACAATCTGAGTAAATGAAATAAATCAG | 1560 |
| Db | 1497 | CTGATTAGTCGATTTACCGCTTTTATAATTTTACAATCTGAGTAAATGAAATAAATCAG | 1556 |
| Qy | 1561 | TTATCTGAAAGCAAATAATCTTTCTGTAACACAGCGTTCGGTCAAAATGGGAAGTTTCATG | 1620 |
| Db | 1557 | TTATCTGAAAGCAAATAATCTTTCTGTAACACAGCGTTCGGTCAAAATGGGAAGTTTCATG | 1616 |
| Qy | 1621 | TGTAATTCATAGTTTATAATAAAGTAAATTTTAAATTTAATTTGTTTATTTTGTTCAGA | 1680 |
| Db | 1617 | TGTAATTCATAGTTTATAATAAAGTAAATTTTAAATTTAATTTGTTTATTTTGTTCAGA | 1676 |
| Qy | 1681 | AATTTAAAAATAAATTTTGTAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACT | 1740 |
| Db | 1677 | AATTTAAAAATAAATTTTGTAGCATGGGAAGTTTACGGGCATCATTTGAGCAGCACTAGACT | 1736 |
| Qy | 1741 | GTTTGAACAATGTATGTCGGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT | 1800 |
| Db | 1737 | GTTTGAACAATGTATGTCGGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT | 1796 |
| Qy | 1801 | GCATTTCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG | 1860 |
| Db | 1797 | GCATTTCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTCAACG | 1856 |

| | | | |
|--|------|--|------|
| Qy | 1861 | GATTGGAAATCCTTTTCTAAACTTTTAAATAAAAAAATGCATTAATGTGTAATATTTATC | 1921 |
| Db | 1857 | GATTGGAAATCCTTTTCTAAACTTTTAAATAAAAAAATGCATTAATGTGTAATATTTATC | 1916 |
| Qy | 1921 | AACACCTCACATTTGATGTTTAGCGTACTATAAATAGTGCTCTTGTGCTCTACTATCAT | 1980 |
| Db | 1917 | AACACCTCACATTTGATGTTTAGCGTACTATAAATAGTGCTCTTGTGCTCTACTATCAT | 1976 |
| Qy | 1981 | CACATCAATCTTACACCACCAAACTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC | 2040 |
| Db | 1977 | CACATCAATCTTACACCACCAAACTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAAC | 2036 |
| Qy | 2041 | ATTCTAAATATC 2052 | |
| Db | 2037 | ATTCTAAATATC 2048 | |
| RESULT 5 | | | |
| LOCUS E40089 | | | |
| DEFINITION Plant promoter and terminator. | | | |
| ACCESSION E40089 | | | |
| VERSION E40089.1 GI:18627205 | | | |
| KEYWORDS JP 2000166577-A/3. | | | |
| SOURCE unidentified | | | |
| ORGANISM unidentified | | | |
| REFERENCE 1 (bases 1 to 2048) | | | |
| AUTHORS Nishikawa,S. and Oeda,K. | | | |
| TITLE Plant promoter and terminator | | | |
| JOURNAL Patent: JP 2000166577-A 3 20-JUN-2000; | | | |
| SUMITOMO CHEM CO LTD | | | |
| COMMENT OS Daucus carota L. | | | |
| PN JP 2000166577-A/3 | | | |
| PD 20-JUN-2000 | | | |
| PF 01-OCT-1999 JP 1999281475 | | | |
| PR | | | |
| PI SATOMI NISHIKAWA, KENJI OEDA | | | |
| PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC C12N15/00, | | | |
| PC C12N5/00,(C12N5/00,C12R1:91) | | | |
| CC | | | |
| FH Key Location/Qualifiers | | | |
| FT promoter (1)..(2048). | | | |
| 1..2048 Location/Qualifiers | | | |
| /organism="unidentified" | | | |
| /mol_type="genomic DNA" | | | |
| /db_xref="taxon:32644" | | | |
| FEATURES | | | |
| source | | | |
| ORIGIN | | | |
| Query Match 99.0%; Score 2032.4; DB 6; Length 2048; | | | |
| Best Local Similarity 99.8%; Pred. No. 3.7e-307; | | | |
| Matches 2047; Conservative 0; Mismatches 1; Indels 4; Gaps 1; | | | |
| Qy | 1 | CATGTGCGCCCTACAGCACATAGGGCTGTTTGGTTGAGAGAAGCAGAGCTGCTTCTGA | 60 |
| Db | 1 | CATGTGCGCCCTACAGCACATAGGGCTGTTTGGTTGAGAGAAGCAGAGCTGCTTCTGA | 60 |
| Qy | 61 | CTTCTCTCTTTTGACCTGTTGTATAAAGAAGTAGAAATATTTTAAAAAGCTGCGAA | 120 |
| Db | 61 | CTTCTCTCTTTTGACCTGTTGTATAAAGAAGTAGAAATATTTTAAAAAGCTGCGAA | 120 |
| Qy | 121 | TACTAACTTCTCTCACAACCTTCGGTCTTTTCCAAACACTTATTAACTTTTACT | 180 |
| Db | 121 | TACTAACTTCTCTCACAACCTTCGGTCTTTTCCAAACACTTATTAACTTTTACT | 180 |
| Qy | 181 | TCTCATTTCTACTCCACTCTTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA | 240 |
| Db | 181 | TCTCATTTCTACTCCACTCTTTTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCA | 240 |
| Qy | 241 | AACGGCCTCAATAAAGATCATTCATAAATGTATCTTTCAATTTTAGGATAACAATACGT | 300 |

Db 241 AACGGCTCAATAAAAGATCAITTCATAAAATGTATCTTTCAAATTTAGGATAACAATACGT 300
Qy 301 GAACAGGTTATTTTAAAGTGTCAACAAATTTCTAAATAATTTTACCTGCGCGGTCAACA 360
Db 301 GAACAGGTTATTTTAAAGTGTCAACAAATTTCTAAATAATTTTACCTGCGCGGTCAACA 360
Qy 361 CCGTCTTCCAAGATATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAAAATTCGCATGC 420
Db 361 CCGTCTTCCAAGATATATATTTTAAATTTTGTAGCTCCCTTTTAAACCAAAATTCGCATGC 420
Qy 421 AGGACGACTAGGTGAATACATTTGTACTGTGAGTCTTTTAAACCAAAATTCGCATGC 480
Db 421 AGGACGACTAGGTGAATACATTTGTACTGTGAGTCTTTTAAACCAAAATTCGCATGC 480
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCGACACACACTCTATCCACGCTACTATCTTTT 540
Db 481 ATGCTCAGCCATCAAAATTTGACAAACCGACACACACTCTATCCACGCTACTATCTTTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAGGATAAGT 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGTAAATTAATGCCCATCCAGGATAAGT 600
Qy 601 AAAATTCGGTTAAACGAGTGTAAATATATATATGTAAATTAATGCCCATCCAGGATAAGT 660
Db 601 AAAATTCGGTTAAACGAGTGTAAATATATATATGTAAATTAATGCCCATCCAGGATAAGT 660
Qy 661 AATACCTTTAGACGACAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCCTAGACT 720
Db 661 AATACCTTTAGACGACAGAGACTTAGGTCAAAATGAGCGCTGGTAAACAGCCTAGACT 720
Qy 721 TGGTCACTGATAAATAGATAAATGTTAGTATATATAGTAGGATCTACAAATGACATTA 780
Db 721 TGGTCACTGATAAATAGATAAATGTTAGTATATATAGTAGGATCTACAAATGACATTA 780
Qy 781 ATTAGAGCTATTAAATGTTACTTAATAAATAGAGAGGTTAGTAAACAGAGAGGTA 840
Db 781 ATTAGAGCTATTAAATGTTACTTAATAAATAGAGAGGTTAGTAAACAGAGAGGTA 840
Qy 841 AAAACAGAGCTTGTCTGT 900
Db 841 AAAACAGAGCTTGTCTGT 900
Qy 901 TAAACTGATCTAAGACACATAGAAATTTAGTACAGGTGTAAGGATCTTAAAGAAATTTATA 960
Db 901 TAAACTGATCTAAGACACATAGAAATTTAGTACAGGTGTAAGGATCTTAAAGAAATTTATA 960
Qy 961 TTAACAGAAATCATTTTATAACATGCTCTCGGTGTGCTATTAATAGGATCACTTAC 1020
Db 961 TTAACAGAAATCATTTTATAACATGCTCTCGGTGTGCTATTAATAGGATCACTTAC 1020
Qy 997 TTAACAGAAATCATTTTATAACATGCTCTCGGTGTGCTATTAATAGGATCACTTAC 1016
Db 997 TTAACAGAAATCATTTTATAACATGCTCTCGGTGTGCTATTAATAGGATCACTTAC 1016
Qy 1021 TGAATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATATCTTACAAATGAA 1080
Db 1021 TGAATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATATCTTACAAATGAA 1080
Qy 1076 TGAATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATATCTTACAAATGAA 1076
Db 1076 TGAATCATCCATTAACACCTTGTAAACAAATTCATGAGATAAATATCTTACAAATGAA 1076
Qy 1081 AAGAAGGACAAATGCTCTTTTGAAGAAACAAATAGGTACTCCCTCGGTCTCTGAAATGT 1140
Db 1081 AAGAAGGACAAATGCTCTTTTGAAGAAACAAATAGGTACTCCCTCGGTCTCTGAAATGT 1140
Qy 1077 AAGAAGGACAAATGCTCTTTTGAAGAAACAAATAGGTACTCCCTCGGTCTCTGAAATGT 1136
Db 1077 AAGAAGGACAAATGCTCTTTTGAAGAAACAAATAGGTACTCCCTCGGTCTCTGAAATGT 1136
Qy 1141 ATACATATCGATTTGACACGGAGACTAAGAAATATGATAAATAGTATAGTATAGTAAAG 1200
Db 1141 ATACATATCGATTTGACACGGAGACTAAGAAATATGATAAATAGTATAGTATAGTAAAG 1200
Qy 1137 ATACATATCGATTTGACACGGAGACTAAGAAATATGATAAATAGTATAGTATAGTAAAG 1196
Db 1137 ATACATATCGATTTGACACGGAGACTAAGAAATATGATAAATAGTATAGTATAGTAAAG 1196
Qy 1201 AAGAAGGAAAGGAGGTTAGTAAAGTACGGAGCCACCAATATATTAATGATAGTATAG 1260
Db 1201 AAGAAGGAAAGGAGGTTAGTAAAGTACGGAGCCACCAATATATTAATGATAGTATAG 1260
Qy 1197 AAGAAGGAAAGGAGGTTAGTAAAGTACGGAGCCACCAATATATTAATGATAGTATAG 1256
Db 1197 AAGAAGGAAAGGAGGTTAGTAAAGTACGGAGCCACCAATATATTAATGATAGTATAG 1256
Qy 1261 AAAAGTAGTTGAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1320
Db 1261 AAAAGTAGTTGAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1320
Qy 1257 AAAAGTAGTTGAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1316
Db 1257 AAAAGTAGTTGAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 1316
Qy 1321 AGAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT 1380
Db 1321 AGAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT 1380
Qy 1317 AGAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT 1376
Db 1317 AGAAGTTTGAATGTATAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAAT 1376

Qy 1381 TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAAATTTTGTATTATTTGATTTC 1440
Db 1377 TAAATGGGACAGAGGAGTAATACCTTTATGATATATAAAATTTTGTATTATTTGATTTC 1436
Qy 1441 TAAAGATTATAAATCTATGTATAAATGATAAATAAATTTTAAATAAATACTATATTAAT 1500
Db 1437 TAAAGATTATAAATCTATGTATAAATGATAAATAAATTTTAAATAAATACTATATTAAT 1496
Qy 1501 CTGATTAGTCGATTACCGCTTTTATATAATTTTCAATCTGAGTAATATGAATAAATCAG 1560
Db 1497 CTGATTAGTCGATTACCGCTTTTATATAATTTTCAATCTGAGTAATATGAATAAATCAG 1556
Qy 1561 TTATCTGAAAAGCAAAATAATATCTTTGTAAGGAGGTTGCGGTCAAAATGCGGAAGTTCATG 1620
Db 1557 TTATCTGAAAAGCAAAATAATATCTTTGTAAGGAGGTTGCGGTCAAAATGCGGAAGTTCATG 1616
Qy 1621 TGTATTCATATAGTTTAAATATAAAGTAAATTTTAAATTTTAAATTTTGTATTTTGTTCAGA 1680
Db 1617 TGTATTCATATAGTTTAAATATAAAGTAAATTTTAAATTTTAAATTTTGTATTTTGTTCAGA 1676
Qy 1681 AATTTAAATAAATTTTGAAGCATCGGAGTTTCAAGGATCATTTGAGCAGCAGCTAGACT 1740
Db 1677 AATTTAAATAAATTTTGAAGCATCGGAGTTTCAAGGATCATTTGAGCAGCAGCTAGACT 1736
Qy 1741 GTTTCGAAACAATCTATGTCGGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1800
Db 1737 GTTTCGAAACAATCTATGTCGGGTGTACATCTATGACCTTTTCAACTCAAACTAGTGAATAAT 1796
Qy 1801 GCATCTTGAATACATCTTTTCAAAATTTTCAAAACACACAGCTTTTAACTTTTCTTCAACG 1860
Db 1797 GCATCTTGAATACATCTTTTCAAAATTTTCAAAACACACAGCTTTTAACTTTTCTTCAACG 1856
Qy 1861 GATTCGAATCTTTTCTAAACTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1920
Db 1857 GATTCGAATCTTTTCTAAACTTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1916
Qy 1921 AACACCTCAACATTTGATGTTAGCTACTATAAATAGGTGCTCTTGTGTGCTCTACTATCAT 1980
Db 1917 AACACCTCAACATTTGATGTTAGCTACTATAAATAGGTGCTCTTGTGTGCTCTACTATCAT 1976
Qy 1981 CACATCAATCTTACACCAAAACCTTGAAGTTTAAATTTTCTACTTATTTCTGAGCAATAAC 2040
Db 1977 CACATCAATCTTACACCAAAACCTTGAAGTTTAAATTTTCTACTTATTTCTGAGCAATCAC 2036
Qy 2041 ATTCTAAATATC 2052
Db 2037 ATTCTAAATATC 2048

RESULT 6

AR076817 LOCUS 2042 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959176.
ACCESSION AR076817
VERSION AR076817.1 GI:10003563
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai, S. and Oeda, K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 2 28-SEP-1999;
FEATURES Location/Qualifiers
1..2042
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 11.8%; Score 241.2; DB 6; Length 2042;
Best Local Similarity 72.8%; Pred. No. 3.5e-28;
Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;

QY 1554 AAATCAGTTATCTGAAAAGCAAATAATATCTTTGTGTAACAACAGCG-----TTGGTCAAAATG 1609
Db 1543 AAATATTATCTGAATGATATCAATCTTTGTGTAACAACAACTGGGCCAAATAGGACCAATAA 1602
QY 1610 GGAAGTTCATGTGTAATCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCAAGTTCACGTGTATCTAAATGTTAATATACTAAACATGAGTATTTCTT--TTCAAGGT 1660
QY 1670 TTTGTTTTCAGAAATTTAAATTAATTTAGCATGGGAAGTTTCAGGGCATCATTTGAGC 1729
Db 1661 ATAAGTTAAATCTTCAATCAATTAATTTAAATTTGGACATTTATTTGACCAACTTTATGCC 1720
QY 1730 AGCACTAGACTGTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAAC 1789
Db 1721 CACGTTGATTTGTTTAAACAACGTTTGTCCGGTGATATTTATGACCTTTCAACTCAAGC 1780
QY 1790 TAGTGAAT-AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAACACAGCACTAACT 1840
QY 1849 TTTCTTTCAACGGAATGGAATCTTTTCTAACTTTTAAATTAATAAAAAATGCAATTAT 1908
Db 1841 TTTCTTTTAAACAGATTAGATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACT 1898
QY 1909 GTAATATTATCAACACCTCAACATGATGTTAGCGTACTATATAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTATCAACACCTCAACATTCATGTTAGCGTACTATATAATAGGTGCTCTTGGTG 1958
QY 1969 CTCTACTATCATCATCAATCTTACACCAAACTTTGAGCTTAATTTTCTACTTATT 2028
Db 1959 CTCTACTATCATCATCAATCTTCCAGCACAACTTTGAGCTTAATCTTCTACTAAT 2018
QY 2029 CTCAGCAATAACATTCTAAATATC 2052
Db 2019 TTTAGCAAAAACATTTCTAAAGGTC 2042

RESULT 7
E15125 LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,T. and Ohta,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/2
PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1..2042 /organism='Daucus carota L.'
FT promoter 1<..2042.
FT Location/Qualifiers
1..2042 /organism='unidentified'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Query Match 11.8%; Score 241.2; DB 6; Length 2042;
Best Local Similarity 72.8%; Pred. No. 3.5e-28;
Matches 367; Conservative 0; Mismatches 128; Indels 9; Gaps 4;
QY 1554 AAATCAGTTATCTGAAAAGCAAATAATATCTTTGTGTAACAACAGCG-----TTGGTCAAAATG 1609
Db 1543 AAATATTATCTGAATGATATCAATCTTTGTGTAACAACAACTGGGCCAAATAGGACCAATAA 1602
QY 1610 GGAAGTTCATGTGTAATCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTATT 1669
Db 1603 CCAAGTTCACGTGTATCTAAATGTTAATATACTAAACATGAGTATTTCTT--TTCAAGGT 1660
QY 1670 TTTGTTTTCAGAAATTTAAATTAATTTAGCATGGGAAGTTTCAGGGCATCATTTGAGC 1729
Db 1661 ATAAGTTAAATCTTCAATCAATTAATTTAAATTTGGACATTTATTTGACCAACTTTATGCC 1720
QY 1730 AGCACTAGACTGTTTGAACAATGATGTCGGGTGATCATCTATGACCTTTCAACTCAAAC 1789
Db 1721 CACGTTGATTTGTTTAAACAACGTTTGTCCGGTGATATTTATGACCTTTCAACTCAAGC 1780
QY 1790 TAGTGAAT-AATGCAATCTAGAAATACATCTTTTCAAAATTTCAACAACACAGCTTTAACT 1848
Db 1781 TAGCCAGTGAATGCTTTCTAGAAATATATCTTTTGAATTTTCAACAACACAGCACTAACT 1840
QY 1849 TTTCTTTCAACGGAATGGAATCTTTTCTAACTTTTAAATTAATAAAAAATGCAATTAT 1908
Db 1841 TTTCTTTTAAACAGATTAGATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACT 1898
QY 1909 GTAATATTATCAACACCTCAACATGATGTTAGCGTACTATATAATAGGTGCTCTTGGTG 1968
Db 1899 ATAATATTATCAACACCTCAACATTCATGTTAGCGTACTATATAATAGGTGCTCTTGGTG 1958
QY 1969 CTCTACTATCATCATCAATCTTACACCAAACTTTGAGCTTAATTTTCTACTTATT 2028
Db 1959 CTCTACTATCATCATCAATCTTCCAGCACAACTTTGAGCTTAATCTTCTACTAAT 2018
QY 2029 CTCAGCAATAACATTCTAAATATC 2052
Db 2019 TTTAGCAAAAACATTTCTAAAGGTC 2042
RESULT 8
AR076816 LOCUS AR076816 247 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 1 from patent US 5959176.
ACCESSION AR076816
VERSION AR076816.1 GI:10003562
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 247)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 1 28-SEP-1999;
FEATURES location/Qualifiers
source 1..247 /organism='unknown'
/mol_type='unassigned DNA'
ORIGIN
Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.7e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1804 TTCTAGATAATCATCTTTTCAAAATTTCAACAACACAGCTTTAACTTTTCTTCAACGGAT 1863
Db 1 TTCTAGAAATATATCTTTTGAATTTTCAACAACACAGCACTAACTTTTCTTTTAAACAGAT 60
QY 1864 TGGAAATCTTTTCTAAACTTTTAAATTAATAAAAAATGCAATTATTTGTAATATTTATCAAC 1923

Db 61 TAGAATCGTTCTCTAAACCTTTTAAAT--AAAAATACATTACTATAATTTATCAAC 118
QY 1924 ACCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1983
Db 119 ACCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1984 ATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAACATT 2043
Db 179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238
QY 2044 CTAATATATC 2052
Db 239 CTAAGGTC 247

RESULT 9
E15124
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 247)
Torikai, T. and Oita, K.,
VEGETABLE PROMOTER AND ITS USE
Patent: JP 1998052273-A 1 24-FEB-1998;
SUMITOMO CHEM CO LTD
OS Dausus carota L. (carrot)
PN JP 1998052273-A/1
PD 24-FEB-1998
PP 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH source
FT
FT
FT
FT
FT

1. .247
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.7e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1804 TTCTAGAATACATCTTTTCAATTTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1863
Db 1 TTCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTAACTTTTCTTTTAAACAGAT 60
QY 1864 TGAATCCCTTTTCTAACTTTTAAATATAAAAAATGATTTTCTAATATTATTTATCAAC 1923
Db 61 TAGAATCGTTCTCTAACTTTTAAATTT--AAAAATACATTACTATAATTTATCAAC 118
QY 1924 ACCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1983
Db 119 ACCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1984 ATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAACATT 2043
Db 179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238

1. .247
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

Query Match 9.6%; Score 196.6; DB 6; Length 247;
Best Local Similarity 89.6%; Pred. No. 4.7e-21;
Matches 223; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1804 TTCTAGAATACATCTTTTCAATTTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1863
Db 1 TTCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTAACTTTTCTTTTAAACAGAT 60
QY 1864 TGAATCCCTTTTCTAACTTTTAAATATAAAAAATGATTTTCTAATATTATTTATCAAC 1923
Db 61 TAGAATCGTTCTCTAACTTTTAAATTT--AAAAATACATTACTATAATTTATCAAC 118
QY 1924 ACCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1983
Db 119 ACCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1984 ATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATAACATT 2043
Db 179 ATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238

QY 2044 CTAATATATC 2052
Db 239 CTAAGGTC 247
RESULT 10
AR146852
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

1 (bases 1 to 246)
Ishige, F., Nishikawa, S. and Oeda, K.,
Plant promoter
Patent: US 6218598-A 2 17-APR-2001;
Location/Qualifiers
1..246
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 6.7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1805 TCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1864
Db 1 TCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTAACTTTTCTTTTAAACAGATT 60
QY 1865 GGAATCCCTTTTCTAACTTTTAAATATAAAAAATGCAATTTGTAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAACA 118
QY 1925 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1984
Db 119 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1985 TCAATCTTACACCAACACCTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAATAACATT 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238
QY 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 11
E55065
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 246)
Ishige, I., Nishikawa, S. and Oeda, K.,
Plant promoter
Patent: JP 2000083679-A 2 28-MAR-2000;
SUMITOMO CHEM CO LTD
OS Dausus carota L.
PN JP 2000083679-A/2
PD 28-MAR-2000
PP 12-JUL-1999 JP 1999197240
PR IKUHARU ISHIGE, SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N15/09,C12R1:91), PC
C12N5/10,C12R1:91), C12N15/00,C12N5/00,C12N15/00,C12R1:91), PC

Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 6.7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1805 TCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1864
Db 1 TCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTAACTTTTCTTTTAAACAGATT 60
QY 1865 GGAATCCCTTTTCTAACTTTTAAATATAAAAAATGCAATTTGTAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAACA 118
QY 1925 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1984
Db 119 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1985 TCAATCTTACACCAACACCTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAATAACATT 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238
QY 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

1. .246
/organism="unidentified"
/mol_type="unassigned DNA"

Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 6.7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1805 TCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1864
Db 1 TCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTAACTTTTCTTTTAAACAGATT 60
QY 1865 GGAATCCCTTTTCTAACTTTTAAATATAAAAAATGCAATTTGTAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAACA 118
QY 1925 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1984
Db 119 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1985 TCAATCTTACACCAACACCTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAATAACATT 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238
QY 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

1. .246
/organism="unidentified"
/mol_type="unassigned DNA"

Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 6.7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;
QY 1805 TCTAGAATACATCTTTTCAAAATTTCAACAAACACAGCTTTAACTTTTCTTCAACGGAT 1864
Db 1 TCTAGAATATATCTTTTGAATTTTCAACAAACACAGCAGCTAACTTTTCTTTTAAACAGATT 60
QY 1865 GGAATCCCTTTTCTAACTTTTAAATATAAAAAATGCAATTTGTAATATTATCAACA 1924
Db 61 AGAATCGTTTCTTAACTTTTAAATTT--AAAAATACATTACTATAATATTATCAACA 118
QY 1925 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 1984
Db 119 CCTCAACATTGATGTAGGTACTATAAATAGTGTCTCTTGGTGTCTTACTATCATCAC 178
QY 1985 TCAATCTTACACCAACACCTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAATAACATT 2044
Db 179 TCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTGTAGCAAAAACATT 238
QY 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

```
(C12N5/00.C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1..246 /organism='Daucus carota L.'
FT Location/Qualifiers
source 1..246
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 9.5%; Score 195.6; DB 6; Length 246;
Best Local Similarity 89.5%; Pred. No. 6.7e-21;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

QY 1805 TCTAGATATATCTTTTGAATTTTCAACACACAGCAGTAACTTTCTTTTCAACGGATT 1864
Db 1 TCTAGATATATCTTTTGAATTTTCAACACACAGCAGTAACTTTCTTTTCAACAGATT 60

QY 1865 GGAATCTCTTTCTAAACTTTTAAATAAAAAATGCAATTAATTAATTTATCAACA 1924
Db 61 AGAATCGTTTCTAAACTTTTAAATTTTAAATAATGCAATTAATTTATCAACA 118

QY 1925 CCTCAACATTGATGTAGGTACTATAATAGGTGCTCTTGGTCTCTACTATCATCA 1984
Db 119 CCTCAACATTGATGTAGGTACTATAATAGGTGCTCTTGGTCTCTACTATCATCA 178

QY 1985 TCAATCTTTACACACAAACCTTGAGCTTAATTTTCTACTTATTTCTCAGCAATTAACATTC 2044
Db 179 TCAATCTTTACACACAAACCTTGAGCTTAATTTTCTACTTATTTTCTCAGCAATTAACATTC 238

QY 2045 TAAATATC 2052
Db 239 TAAAGGTC 246

RESULT 12
DCA18706 DCA18706 8005 bp DNA linear PLN 10-FEB-1999
LOCUS Daucus carota Inv*Dc5 gene.
DEFINITION Daucus carota Inv*Dc5 gene.
ACCESSION Y18706
VERSION Y18706.1 GI:4454114
KEYWORDS beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform.
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae;
Daucinae; Daucus.
1
REFERENCE
AUTHORS Sturm,A.
TITLE Molecular characterisation and functional analysis of
sucrose-cleaving enzymes in carrot (Daucus carota L.)
JOURNAL J. Exp. Bot. 47, 1187-1192 (1996)
REFERENCE
AUTHORS Sturm,A.
TITLE Direct Submission
JOURNAL Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut,
Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND
COMMENT Related sequence X67163.
FEATURES
source 1..8005
/organism="Daucus carota"
/mol_type="genomic DNA"
/cultivar="Nantaise"
/db_xref="taxon:4039"
1..8005
/gene="Inv*Dc5"
1..2666
/gene="Inv*Dc5"
Join(2767..3160,3724..3732,4656..5515,5605..5766,
5869..6104,6190..6277,6384..6587)
CDS
5869..6104,6190..6277,6384..6587)

/gene="Inv*Dc5"
/EC number="3.2.1.26"
/notes="soluble acid invertase"
/codon_start=1
/product="beta-fructofuranosidase, isoform II"
/protein_id="CAA77266.1"
/db_xref="GI:4454115"
/db_xref="GOA:Q42722"
/db_xref="TrEMBL:Q42722"
/translation="MEHPITISHVTPLPDGHSPSLATTNTAEQSSRRSLTFVLLFS
SILAACLVMGTMLFPNSGNEAVEKSTVPEETVEVAPRGVAEGVSMKSPRRPALNAE
PPANFPNSNVLSHQSRSPHPQPNQNMNDPGLFYKGWYHLFYQYNPDGAINGNKI
VMHAYSSDLIHWKHLFPAMVTDHWYDVGWVTSATLPDQIIVMLTIGSTNESVQV
QNLAYADPDDPLLIIEWKYPGNVLPVPPGIDFKOPDPTTAWRTPEGRWLTIGSK
LNKTGILSVDTVDNFYTLDDGLVHVAHGTGMWECVDFYVSKFGENGLTISFDVG
VHKYMASLDDDRNDYVAIGTYDPSGKVPDNPDELVDIGLRDYDGIYYASKTFYDS
NKKRRVLSMKETDSEISDVVRKQWASVQGIPTLILPDKTGSNLLQWPEVKNLRL
NKTYPENVEINTGAVLPLEIGSGSOLDITAEFEVDKSLERVOETNEVYDCKNNGKLS
GRGALGFGLLILADKDLSEQTPIVYTIAGSGGNLRTFFCADHSRSSKAVDVKETIY
GSVPVLURGSKLTMRIILVDHSIVESFSQGGRTCTISRYVPTKAIYNNNAKVLFLFNATE
ARITASLNIWQMNTAQRQTHPADLVI"
2767..3160
/gene="Inv*Dc5"
/number=1
3161..3723
/gene="Inv*Dc5"
/number=1
3724..3732
/gene="Inv*Dc5"
/number=2
3733..4655
/gene="Inv*Dc5"
/number=2
4656..5515
/gene="Inv*Dc5"
/number=3
5516..5604
/gene="Inv*Dc5"
/number=3
5605..5766
/gene="Inv*Dc5"
/number=4
5767..5868
/gene="Inv*Dc5"
/number=4
5869..6104
/gene="Inv*Dc5"
/number=5
6105..6199
/gene="Inv*Dc5"
/number=5
6190..6277
/gene="Inv*Dc5"
/number=6
6278..6383
/gene="Inv*Dc5"
/number=6
6384..6587
/gene="Inv*Dc5"
/number=7
6588..8005
/gene="Inv*Dc5"

Query Match 8.5%; Score 174; DB 8; Length 8005;
Best Local Similarity 81.7%; Pred. No. 7.8e-18;
Matches 201; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 17 CACATAGGGCCTGTTTGGTTGAGAGACGACAGCTGCTTCTGACTTCTTCTTTGA 76
Db 1285 CACTATGGGGTGTGTTGTTGTTTCCCGAAGTAGAGCTGCTTCGTTCTTCTTTGA 1226
QY 77 CCTGTTTGTATATAAGAGTAGAATAATATTTTTAAAAAGCTGCGAATACTTCTCTC 136
```

Db 1225 CCGTGTGTAATAAGCAGAGCAATTTTAAAAAGTTGAGAATGCTAGCTTCTCTCTC 1166
QY 137 AGCACTTCGCTTCTTTTCCAAACACTTATTAATCTTTTACTTCTCAATTTCTACTCCA 196
Db 1165 ACAGCTTCGCTTCTTTTCCAAACACTTATTAATCTTTTACTTCTCACTTCTACTCTA 1106
QY 197 CTTCTTTGCTATAAGCAAGAAATCACTCTTTTAAAGCTAACCAACGGCCTCAATAAAA 256
Db 1105 CTTTCTTAATTTAAGTAAGAAATCACTCTTTTAAAGCTAACCAACGGCCTCAATGACT 1046
QY 257 GATCAT 262
Db 1045 GACCT 1040

RESULT 13
BD188672
LOCUS
DEFINITION Promotor and use thereof. 2831 bp DNA linear PAT 17-JUL-2003
ACCESSION BD188672
VERSION BD188672.1 GI:32998411
KEYWORDS JP 2003000252-A/1.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2831)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof
JOURNAL Patent: JP 2003000252-A 1 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/1
PD 07-JAN-2003
PP 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/10
PC 00,C12N5/00
CC Promotor and use thereof
FH Key Location/Qualifiers
FT Promoter (1)..(2831).
FEATURES
source
1..2831
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2831;
Best Local Similarity 82.5%; Pred. No. 6.1e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 22 AGGGCTGTTGGTTGAGAGAGCAGAGCTGTTCTGACTTCTTCTTCTTTTGACCTGT 81
Db 1716 AGGGCTGTTGGTTTATGGAATCAGAAGCTGTTCTGACTTCTTCTTCTTCTTCT 1774
QY 82 TTGTATAAGAGTAGAATAATATTTTAAAAAGCTCGAATACTAACTTCTCTCTCAAC 141
Db 1775 TTGTGTAATAAGAGCAGAGCACTTTTAAAGAGCTGAGAATGCTAGTCTCTCTCACAGC 1834
QY 142 TTCGGCTTCTTTTCCAAACACTTATTAATCTTTTACTTCTCATTTCTTCTTCTTCT 201
Db 1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTTCTTCTTCTTCTTCTTCTTCT 1894
QY 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAACCCAAAGCGCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAGTCAATCTTTTAAATTAACCCAAAGCGCCCTTAAGTAATTGA 1954

RESULT 14
BD188674
LOCUS
DEFINITION Promotor and use thereof. 2865 bp DNA linear PAT 17-JUL-2003
ACCESSION BD188674
VERSION BD188674.1 GI:32998413
KEYWORDS JP 2003000252-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2865)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof
JOURNAL Patent: JP 2003000252-A 3 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/3
PD 07-JAN-2003
PP 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/10
PC 00,C12N5/00
CC Promotor and use thereof
FH Key Location/Qualifiers
FT Promoter (1)..(2865).
FEATURES
source
1..2865
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2831;
Best Local Similarity 82.5%; Pred. No. 6.1e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 22 AGGGCTGTTGGTTGAGAGAGCAGAGCTGTTCTGACTTCTTCTTCTTTTGACCTGT 81
Db 1716 AGGGCTGTTGGTTTATGGAATCAGAAGCTGTTCTGACTTCTTCTTCTTCTTCT 1774
QY 82 TTGTATAAGAGTAGAATAATATTTTAAAAAGCTCGAATACTAACTTCTCTCTCAAC 141
Db 1775 TTGTGTAATAAGAGCAGAGCACTTTTAAAGAGCTGAGAATGCTAGTCTCTCTCACAGC 1834
QY 142 TTCGGCTTCTTTTCCAAACACTTATTAATCTTTTACTTCTCATTTCTTCTTCTTCTTCT 201
Db 1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTTCTTCTTCTTCTTCTTCTTCT 1894
QY 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAACCCAAAGCGCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAGTCAATCTTTTAAATTAACCCAAAGCGCCCTTAAGTAATTGA 1954

ACCESSION BD188674
VERSION BD188674.1 GI:32998413
KEYWORDS JP 2003000252-A/3.
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 2865)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Promotor and use thereof
JOURNAL Patent: JP 2003000252-A 3 07-JAN-2003;
SUMITOMO CHEMICAL CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 2003000252-A/3
PD 07-JAN-2003
PP 31-MAY-2001 JP 2001164069
PI SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/15,C12N1/19,C12N1/21,C12N5/10,C12N15/10
PC 00,C12N5/00
CC Promotor and use thereof
FH Key Location/Qualifiers
FT Promoter (1)..(2865).
FEATURES
source
1..2865
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 7.9%; Score 162.4; DB 6; Length 2865;
Best Local Similarity 82.5%; Pred. No. 6.1e-16;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 22 AGGGCTGTTGGTTGAGAGAGCAGAGCTGTTCTGACTTCTTCTTCTTTTGACCTGT 81
Db 1716 AGGGCTGTTGGTTTATGGAATCAGAAGCTGTTCTGACTTCTTCTTCTTCTTCT 1774
QY 82 TTGTATAAGAGTAGAATAATATTTTAAAAAGCTCGAATACTAACTTCTCTCTCAAC 141
Db 1775 TTGTGTAATAAGAGCAGAGCACTTTTAAAGAGCTGAGAATGCTAGTCTCTCTCACAGC 1834
QY 142 TTCGGCTTCTTTTCCAAACACTTATTAATCTTTTACTTCTCATTTCTTCTTCTTCT 201
Db 1835 TTCTGCTTCTTTTCCAAACACTTTATCAACTTACCTTCTTCTTCTTCTTCTTCTTCT 1894
QY 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAAGCTAACCCAAAGCGCTCAATAAAGATCA 261
Db 1895 TTACTATAAGCAAGAGTCAATCTTTTAAATTAACCCAAAGCGCCCTTAAGTAATTGA 1954
RESULT 15
DARGCHS2
LOCUS
DEFINITION Carrot GCHS2 gene for chalcone synthase. 4886 bp DNA linear PLN 14-APR-2000
ACCESSION D16255
VERSION D16255.1 GI:441168
KEYWORDS chalcone synthase.
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae; Daucus.
REFERENCE 1 (bases 1 to 4886)
AUTHORS Ozeki,Y., Davies,E. and Takeda,J.
TITLE Structure and expression of chalcone synthase gene in carrot suspension cultured cells regulated by 2,4-D
JOURNAL Plant Cell Physiol. 34, 1029-1037 (1993)
REFERENCE 2 (bases 1 to 4886)
AUTHORS Ozeki,Y.
TITLE Direct Submission
JOURNAL Submitted (14-MAY-1993) Yoshihiro Ozeki, College of Arts and Sciences, The University of Tokyo, Department of Biology, Komaba, Meguro-ku, Tokyo 153, Japan (Tel:03-3467-1171(ex.253),

Search completed: December 6, 2004, 22:47:27
Job time : 9444.04 secs

COMMENT

Fax: 03-3485-2904)
Submitted (14-MAY-1993) to DDBJ by:
Yoshihiro Ozeki
Department of Biol. College of
Arts & Science, Univ. of Tokyo
3-8-1 Komaba, Meguro-ku
Tokyo 153

| | | |
|--------------|-------------|--|
| intron | 2608. .3394 | GNSSACVLFILDEMRKASAKDGHKIIIGGLDGVGVFGFGGLIVETVVLHVSVP. |
| exon | 3395. .4519 | /gene="gCHS2" |
| polyA signal | 4501. .4506 | /number=2 |

| Query Match | 7.9% | Score 162.4 | DB 8 | Length 4886 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 82.5% | Pred. No. 5.5e-16 | | |
| Matches 198 | Conservative 0 | Mismatches 41 | Indels 1 | Gaps 1 |
| Qy | 11 | CTACAGCACATAGGCCCTGTTTGGTCTCAGAGAGCAGAAAGCTGCTTCGTAGACTTCTTCTTC | 70 | |
| Db | 1355 | CTAGTAGATGAGGCCCTGTTTGGATGGTGGAGCAGAAAGCTGCTTCGGCTTCTGCTTT | 1414 | |
| Qy | 71 | TTTTGACCTGTTTGATATAAGAAAGTAGAATAATTTTTAAAAAGCTCGAAATACTAACTTC | 130 | |
| Db | 1415 | TCTTGACCCGTTTGGGTAAAGAGTAGAAGCAGCTTTTAAAGAAGTTGAGAAATACTAGCTTC | 1474 | |
| Qy | 131 | TCTCTCACAACTTCG-CTTCTTTTCCAAACACTTTATTAACTTTTTTACTTCTCATTTTC | 189 | |
| Db | 1475 | TCCTCAGAGGCTCTGTTTTTTTTTTCCAAACACTTTATTATTATTCTTTTCACTTC | 1534 | |
| Qy | 190 | TACTCCACTTCTTTTGCTATAAGCAAGAAATCACTTCTTTTAAAGTAAACCCAAACGGCCTC | 249 | |
| Db | 1535 | TACTCCACTTCTTTAAATTTAAGCAAGAGTCACTCTTTTAAACTAACTAAACCAACGGCCCC | 1594 | |

This sequence represents a carrot promoter. The invention relates to plant promoters and terminators from *Daucus carota* L. which are capable of expressing a gene of interest in plants. The invention also includes a chimeric gene characterized in that it comprises the promoter and a desired gene linked to each other in the form capable of functioning. A method of producing a transformant comprises introducing the promoter, the chimeric gene or a vector comprising the promoter and a desired gene or terminator sequence into a host cell. The plant promoters and terminators are useful in plant breeding, for e.g. fertilities of plants may be controlled by expressing, in the host cells, a sense or antisense gene of a male sterility related gene such as S-locus-specific RNase gene

XX SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;
Query Match 100.0%; Score 2052; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2052; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CATGTGCGCCCTACAGCACAATAGGCGCTGTTGGTTGAGAGAGAGAGAGCTGCTCTGA 60
DB |||||
QY 1 CATGTGCGCCCTACAGCACAATAGGCGCTGTTGGTTGAGAGAGAGAGAGCTGCTCTGA 60
DB |||||
QY 61 CTTCT 120
DB |||||
QY 61 CTTCT 120
DB |||||
QY 121 TACTAATCT 180
DB |||||
QY 121 TACTAATCT 180
DB |||||
QY 181 TCTCATTCT 240
DB |||||
QY 181 TCTCATTCT 240
DB |||||
QY 241 AACGGCCTCAATAAAGATCATCTAATATGATCTTTCAATTTTGTAGGATACCAATACGT 300
DB |||||
QY 241 AACGGCCTCAATAAAGATCATCTAATATGATCTTTCAATTTTGTAGGATACCAATACGT 300
DB |||||
QY 301 GAACAGGGTTATTTTTTAAACGTTCAACAAATCTTAATAATTTTACCTGGCGGTGAACA 360
DB |||||
QY 301 GAACAGGGTTATTTTTTAAACGTTCAACAAATCTTAATAATTTTACCTGGCGGTGAACA 360
DB |||||
QY 361 CGGTCTTCCAGATAATATATTTTAAATTTTGTAGTCTCTCTCTCTCTCTCTCTCTCTCT 420
DB |||||
QY 361 CGGTCTTCCAGATAATATATTTTAAATTTTGTAGTCTCTCTCTCTCTCTCTCTCTCTCT 420
DB |||||
QY 421 AGGACGACTTAGGTGAATACACATCTGTACTGTAGTCTTTTAAACAAAGAAACAAGTGTTC 480
DB |||||
QY 421 AGGACGACTTAGGTGAATACACATCTGTACTGTAGTCTTTTAAACAAAGAAACAAGTGTTC 480
DB |||||
QY 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATACCTTT 540
DB |||||
QY 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACACACTCTATCCAGTACTATACCTTT 540
DB |||||
QY 541 TGGCCGATGCTCTCAAAATGTTTTTATATATATATATATATATATATATATATATATATAT 600
DB |||||
QY 541 TGGCCGATGCTCTCAAAATGTTTTTATATATATATATATATATATATATATATATATATAT 600
DB |||||
QY 601 AAAATTTCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATATAT 660
DB |||||
QY 601 AAAATTTCCCGTTTAAACAGTTTGTATATATATATATATATATATATATATATATATATAT 660
DB |||||
QY 661 AATATCTTTTACGACAGAGACTTAGGTCAAAAATGAGCGCTGTGTATATATATATATATAT 720
DB |||||
QY 661 AATATCTTTTACGACAGAGACTTAGGTCAAAAATGAGCGCTGTGTATATATATATATATAT 720
DB |||||
QY 721 TGGTCACTGATATAATAGATAATTTGTAGTATATATATATATATATATATATATATATATAT 780
DB |||||
QY 721 TGGTCACTGATATAATAGATAATTTGTAGTATATATATATATATATATATATATATATATAT 780
DB |||||
QY 781 ATTAGAGCTATTAATTAAGTACTAATAAATAGAGAGTTAGTATAACAGAAAGCAGGTA 840
DB |||||
QY 781 ATTAGAGCTATTAATTAAGTACTAATAAATAGAGAGTTAGTATAACAGAAAGCAGGTA 840
DB |||||
QY 841 AAAACAAGAGCTGT 900
DB |||||
QY 841 AAAACAAGAGCTGT 900
DB |||||
QY 901 TAAACTGATCTAAGCAGATAGAAATTTAGTACAGGTTTAAACTTTTACAGAAATTTTATA 960
DB |||||
QY 901 TAAACTGATCTAAGCAGATAGAAATTTAGTACAGGTTTAAACTTTTACAGAAATTTTATA 960
DB |||||
QY 961 TTAAACGAAATCATTTTTTATAACATGTCTCTCGGCTGTCATTATATATAGGGATCACTTAC 1020
DB |||||

DB |||||
QY 961 TTAAACGAAATCATTTTTTATAACATGTCTCTCGGCTGTCATTATAATAGGATCACCTTAC 1020
DB |||||
QY 1021 TGATCATCCATTAACAACTTGTTTAAACAAATCAATGAGATAAAATATCTTACAATGAA 1080
DB |||||
QY 1021 TGATCATCCATTAACAACTTGTTTAAACAAATCAATGAGATAAAATATCTTACAATGAA 1080
DB |||||
QY 1081 AAGAGGACATGTCTCTTTGAAAAAACAATAGGTACTCCCTCGGCTCTGAAATGT 1140
DB |||||
QY 1081 AAGAGGACATGTCTCTTTGAAAAAACAATAGGTACTCCCTCGGCTCTGAAATGT 1140
DB |||||
QY 1141 ATACATATGGATTGGACACGAGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
DB |||||
QY 1141 ATACATATGGATTGGACACGAGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAAAAG 1200
DB |||||
QY 1201 AAGAGAAAGAAAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTTAG 1260
DB |||||
QY 1201 AAGAGAAAGAAAAAGTGGGTAAAGTAGCGGACCCCAATATATAATTTGATAGATTTAG 1260
DB |||||
QY 1261 AAAAGTAGTTCAAAAGTAGTGGTGGGTGGGATTTTTTATATATAAAAAATTTACTATTTTG 1320
DB |||||
QY 1261 AAAAGTAGTTCAAAAGTAGTGGTGGGTGGGATTTTTTATATATAAAAAATTTACTATTTTG 1320
DB |||||
QY 1321 AGAAAGTTTTGAAATGTATAGAAATTTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAA 1380
DB |||||
QY 1321 AGAAAGTTTTGAAATGTATAGAAATTTGAGTGGGACATCCATAAAAGGAAAGTGTATAGAA 1380
DB |||||
QY 1381 TAAATGGGACAGAGGGAGTAATACCTTTATATATATAAATTTTGTATTTGATTTTCA 1440
DB |||||
QY 1381 TAAATGGGACAGAGGGAGTAATACCTTTATATATATAAATTTTGTATTTGATTTTCA 1440
DB |||||
QY 1441 TAAGATTATTAATCTATGTAT 1500
DB |||||
QY 1441 TAAGATTATTAATCTATGTAT 1500
DB |||||
QY 1501 CTGATTAGTCTGATTACCGCTTTTATAATTTTCAATACTGAGTAATATGAATAAATCAG 1560
DB |||||
QY 1501 CTGATTAGTCTGATTACCGCTTTTATAATTTTCAATACTGAGTAATATGAATAAATCAG 1560
DB |||||
QY 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAATGGGAAAGTTCAATG 1620
DB |||||
QY 1561 TTATCTGAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAATGGGAAAGTTCAATG 1620
DB |||||
QY 1621 TGATTTCAATAGTTTAAATATAAAGTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1680
DB |||||
QY 1621 TGATTTCAATAGTTTAAATATAAAGTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAA 1680
DB |||||
QY 1681 AATTTAAAAATAAATTTATTGAGCATGGGAAGTTCAACGGGCATCATTTGAGCAGCAGTACT 1740
DB |||||
QY 1681 AATTTAAAAATAAATTTATTGAGCATGGGAAGTTCAACGGGCATCATTTGAGCAGCAGTACT 1740
DB |||||
QY 1741 GTTTGAAACAATGTATGTCCGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATAT 1800
DB |||||
QY 1741 GTTTGAAACAATGTATGTCCGGTGATCATCTATGACCTTTTCAACTCAAACTAGTGAATAT 1800
DB |||||
QY 1801 GCATTTCTAGATACATCTTTTCAAAATTTCAACAAACAGCGCTTTTAACTTTTCTTCAACG 1860
DB |||||
QY 1801 GCATTTCTAGATACATCTTTTCAAAATTTCAACAAACAGCGCTTTTAACTTTTCTTCAACG 1860
DB |||||
QY 1861 GATTGGAATCCTTTTCTAAACTTTTAAAAATAAAAAATGCAATTTTGTAAATTTTATC 1920
DB |||||
QY 1861 GATTGGAATCCTTTTCTAAACTTTTAAAAATAAAAAATGCAATTTTGTAAATTTTATC 1920
DB |||||
QY 1921 AACACCTCAACATTTGATGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTATCAT 1980
DB |||||
QY 1921 AACACCTCAACATTTGATGTAGCGTACTATAAATAGGTGCTCTTGGTGTCTCTATCAT 1980
DB |||||
QY 1981 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATTAAC 2040
DB |||||
QY 1981 CACATCAATCTTACACCAAAACCTTGAGCTTAATTTTCTTACTTATTTCTCAGCAATTAAC 2040
DB |||||
QY 2041 ATCTCAATATC 2052
DB |||||
QY 2041 ATCTCAATATC 2052
DB |||||

Db 1381 ||||| TAAATGGGACAGGGAGTAACCTTATATGATATATAATTTTGTATTTGATTTCA 1440
Qy 1441 ||||| TAAAGTATATAATCTATGATTAATAATGATATAATAATTTTAAAAATAATACTATATAATT 1500
Db 1441 ||||| TAAAGTATATAATCTATGATTAATAATGATATAATAATTTTAAAAATAATACTATATAATT 1500
Qy 1501 ||||| CTGATTAGTCGATTACCGCTTTTATATTTTACAACTAGTGAATATGATTAATCAATCAG 1560
Db 1501 ||||| CTGATTAGTCGATTACCGCTTTTATATTTTACAACTAGTGAATATGATTAATCAATCAG 1560
Qy 1561 ||||| TTATCTGAAAAGCAATAATATCTTTGTAATAACAGCGTTTCGGTCAAAATGGGAAGTTCAATG 1620
Db 1561 ||||| TTATCTGAAAAGCAATAATATCTTTGTAATAACAGCGTTTCGGTCAAAATGGGAAGTTCAATG 1620
Qy 1621 ||||| TGTATCTCAATAGTTTAAATATATAAAGTAAATTTTAAATTTAAATTTGTTTTCAGA 1680
Db 1621 ||||| TGTATCTCAATAGTTTAAATATATAAAGTAAATTTTAAATTTAAATTTGTTTTCAGA 1680
Qy 1681 ||||| AATTTAAATAAATTTATTGAGCATGGGAAGTTTCACGGGCATCATTTGAGCAGCACTAGACT 1740
Db 1681 ||||| AATTTAAATAAATTTATTGAGCATGGGAAGTTTCACGGGCATCATTTGAGCAGCACTAGACT 1740
Qy 1741 ||||| GTTTGAACAATGTATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Db 1741 ||||| GTTTGAACAATGTATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
Qy 1801 ||||| GCATCTCTAGAAATACATCTTTTCAAAATTTTCAACAAACACAGCTTTTCTTTTCAACG 1860
Db 1801 ||||| GCATCTCTAGAAATACATCTTTTCAAAATTTTCAACAAACACAGCTTTTAACTTTTCTTTCAACG 1860
Qy 1861 ||||| GATTGGAATCCCTTTTCTAAACCTTTTAAATATAAATAAATAATGATTTTGTATATTTATC 1920
Db 1861 ||||| GATTGGAATCCCTTTTCTAAACCTTTTAAATATAAATAAATAATGATTTTGTATATTTATC 1920
Qy 1921 ||||| AACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGCTGCTCTACTATCAT 1980
Db 1921 ||||| AACACCTCAACATTTGATGTTAGCGTACTATAAATAGGTGCTCTTGCTGCTCTACTATCAT 1980
Qy 1981 ||||| CACATCAATCTTACACCAAAACCTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAATAAC 2040
Db 1981 ||||| CACATCAATCTTACACCAAAACCTTGAGCTTTAAATTTTCTACTTATTTCTCAGCAATCAC 2040
Qy 2041 ||||| ATTCTAAATATC 2052
Db 2041 ||||| ATTCTAAAGATC 2052

RESULT 3
AAA37964
ID AAA37964 standard; DNA; 2056 BP.
XX
AC AAA37964;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO2000020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;

XX WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene,
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;
Query Match 99.2%; Score 2034.8; DB 3; Length 2056;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2050; Conservative 0; Mismatches 2; Indels 4; Gaps 1;
Qy 1 CATGTGTGCCCTACAGCACATAGGGCCTGTTGGTTGAGAGAAGCAGAGCTGCTTCTGA 60
Db 1 CATGTGTGCCCTACAGCACATAGGGCCTGTTGGTTGAGAGAAGCAGAGCTGCTTCTGA 60
Qy 61 CTTCTTCTCTCTCTCACAACTTCGGCTCTCTTTCCAAACACTTTATTAACTTTTAACT 120
Db 61 CTTCTTCTCTCTCTCACAACTTCGGCTCTCTTTCCAAACACTTTATTAACTTTTAACT 120
Qy 121 TACTAACTTCTCTCTCACAACTTCGGCTCTCTTTCCAAACACTTTATTAACTTTTAACT 180
Db 121 TACTAACTTCTCTCTCACAACTTCGGCTCTCTTTCCAAACACTTTATTAACTTTTAACT 180
Qy 181 TCTCATTTCTACTCCACTTCTTGTCTATAAGCAAGAAATCACTCTTTTAAAGCTAAACCA 240
Db 181 TCTCATTTCTACTCCACTTCTTGTCTATAAGCAAGAAATCACTCTTTTAAAGCTAAACCA 240
Qy 241 AACGGCCTCAATAAAGATCATTTCAATATGATCTTTCAATTTTAGATTAACATAGCT 300
Db 241 AACGGCCTCAATAAAGATCATTTCAATATGATCTTTCAATTTTAGATTAACATAGCT 300
Qy 301 GAACAGGGTTATTTTAAACGTTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA 360
Db 301 GAACAGGGTTATTTTAAACGTTCAACAAATTTCTAATAATTTTACCTGGCGGTGAACA 360
Qy 361 CCGTCTTCCAAAGATAATATTTTAAATTTGTAGCTCCCTTTTAAACCAAAATTCGCATGC 420
Db 361 CCGTCTTCCAAAGATAATATTTTAAATTTGTAGCTCCCTTTTAAACCAAAATTCGCATGC 420
Qy 421 AGGACGACTTAGGTGAATACATTTGATCTGTGAGTCTTTTAAACAAAGAAACAAGTGGTTC 480
Db 421 AGGACGACTTAGGTGAATACATTTGATCTGTGAGTCTTTTAAACAAAGAAACAAGTGGTTC 480
Qy 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGCTACTATATCTT 540
Db 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCAGCTACTATATCTT 540
Qy 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGATTAATAATATGCCATTCGAAGGATAAGT 600
Db 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATGATTAATAATATGCCATTCGAAGGATAAGT 600
Qy 601 AAAATTCGGTTTAAACGAGTTTGTATATATATATATATATATATATATATATATATAT 660
Db 601 AAAATTCGGTTTAAACGAGTTTGTATATATATATATATATATATATATATATATATAT 660
Qy 661 AATACTTTTACGACGAAGAGACTTAGGTCAAAATGGACGCTGGTAAACAGCCCTAGACT 720
Db 661 AATACTTTTACGACGAAGAGACTTAGGTCAAAATGGACGCTGGTAAACAGCCCTAGACT 720

AC AAA37963;
 XX 18-AUG-2000 (first entry)
 XX Plasmid #2 DNA sequence used in mutation of promoter sequence.
 DE Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
 KW Daucus carota.
 OS WO2000206113-A1.
 XX 13-APR-2000.
 XX 28-SEP-1999; 99WO-JP005303.
 XX 02-OCT-1998; 98JP-00281124.
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX Nishikawa S, Oeda K;
 PI WPI; 2000-303791/26.
 DR New Plant promoters and terminators from Daucus carota L., useful in
 XX plant breeding, for e.g. controlling fertilities of plants.
 XX Example 8; Page 73-74; 81pp; English.
 CC This sequence represents a plasmid sequence used in a method for
 CC introducing a mutation into a carrot promoter. The invention relates to
 CC plant promoters and terminators from Daucus carota L. which are capable
 CC of expressing a gene of interest in plants. The invention also includes a
 CC chimeric gene characterized in that it comprises the promoter and a
 CC desired gene linked to each other in the form capable of functioning. A
 CC method of producing a transformant comprising introducing the promoter,
 CC the chimeric gene or a vector comprising the promoter and a desired gene
 CC or terminator sequence into a host cell. The plant promoters and
 CC terminators are useful in plant breeding, for e.g. fertilities of plants
 CC may be controlled by expressing, in the host cells, a sense or antisense
 CC gene of a male sterility related gene such as S-locus-specific RNase gene
 XX Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;
 SQ

Query Match 99.0%; Score 2030.8; DB 3; Length 2048;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 2046; Conservative 0; Mismatches 2; Indels 4; Gaps 1;

QY 1 CATGTGCGCTACAGCACATAGGGCTCTTTGGTTGAGAGAGAGAGAGCTGCTTCTGA 60
 DB 1 CATGTGCGCTACAGCACATAGGGCTCTTTGGTTGAGAGAGAGAGAGCTGCTTCTGA 60
 QY 61 CTTCTCTCTTTTGACCTGTTGTATAAAGAGTAGAAATATTTTTAAAAAGTCGGAA 120
 DB 61 CTTCTCTCTTTTGACCTGTTGTATAAAGAGTAGAAATATTTTTAAAAAGTCGGAA 120
 QY 121 TACTAACTCTCTCACAACTTCGCTCTTTCCAAACACTTTATTAACCTTTTACT 180
 DB 121 TACTAACTCTCTCACAACTTCGCTCTTTCCAAACACTTTATTAACCTTTTACT 180
 QY 181 TCTCATTTCTCTCTTCTTGTATTAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240
 DB 181 TCTCATTTCTCTCTTCTTGTATTAAGCAAGAAATCACTCTTTTAAGCTAACCCA 240
 QY 241 AACGGCTCAATAAAGATCATTAATGATCTTTCAATTTAGGTAAACAATACGT 300
 DB 241 AACGGCTCAATAAAGATCATTAATGATCTTTCAATTTAGGTAAACAATACGT 300
 QY 301 GAACAGGGTATTTTACAGTGTCAACAAATCTTAATATTTTACCTGGCGGTGAACA 360
 DB 301 GAACAGGGTATTTTACAGTGTCAACAAATCTTAATATTTTACCTGGCGGTGAACA 360
 QY 361 CGGTCTTCCAAGATAATAATTTTAAATTTTGTAGCTCCCTTTTAACCAAAATTCGGATGC 420

DB 361 CGGTCTTCCAAGATAATAATTTTAAATTTTGTAGCTCCCTTTTAACCAAAATTC---GC 416
 QY 421 AGGACACTTAGTGAATACACATTTGTACTGTAGTCTTTAAACAAAGAACAGTGGTTC 480
 DB 417 AGGACACTTAGTGAATACACATTTGTACTGTAGTCTTTAAACAAAGAACAGTGGTTC 476
 QY 481 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTACTATACTTT 540
 DB 477 ATGCTCAGCCATCAAAATTTGACAAACCCGACACAACTCTATCCACGTACTATACTTT 536
 QY 541 TGGCCGAATGCTTCTCAAAATTTTATATGTAATAATAATGCCCATCCAAAGGATAAGT 600
 DB 537 TGGCCGAATGCTTCTCAAAATTTTATATGTAATAATAATGCCCATCCAAAGGATAAGT 596
 QY 601 AAAATTCGCGTTAAACAGTTTGTAAATATATATGTTACACTTACAGAGGATATTCGT 660
 DB 597 AAAATTCGCGTTAAACAGTTTGTAAATATATATGTTACACTTACAGAGGATATTCGT 656
 QY 661 AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGACGCTGGTAAACAGCCTAGACT 720
 DB 657 AATACTTTTAGACGACAGAGACTTAGGTCAAAAATGACGCTGGTAAACAGCCTAGACT 716
 QY 721 TGGTCACTGATAAATAGATAATGTTAGTATAATAATAGTAGGATCTACAATGACATAAA 780
 DB 717 TGGTCACTGATAAATAGATAATGTTAGTATAATAATAGTAGGATCTACAATGACATAAA 776
 QY 781 ATTAGAGCTATTAATTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAACGAGTA 840
 DB 777 ATTAGAGCTATTAATTAAGTTACTATAAATAAGAGAGGTTAGTAAACAGAACGAGTA 836
 QY 841 AAAACAAGAGCTGCTGCTGTTAGTTAGTCTGTAGCTCATTTCTTTAAAGTAATG 900
 DB 837 AAAACAAGAGCTGCTGCTGTTAGTTAGTCTGTAGCTCATTTCTTTAAAGTAATG 896
 QY 901 TAAACTGATCTAAAGCACATGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTA 960
 DB 897 TAAACTGATCTAAAGCACATGAAATTTAGTACAGGTTAAAACTTTTACAAGAAATTA 956
 QY 961 TTAACGAAATCATTTTATACATGCTCTCGGCTGTCATTTATATAGGGATCACTTAC 1020
 DB 957 TTAACGAAATCATTTTATACATGCTCTCGGCTGTCATTTATATAGGGATCACTTAC 1016
 QY 1021 TGATCATCTCAATTAACCTTTGTTAAAAACAAATTCATGAGATAAAATATCTTACAATGAA 1080
 DB 1017 TGATCATCTCAATTAACCTTTGTTAAAAACAAATTCATGAGATAAAATATCTTACAATGAA 1076
 QY 1081 AAGAAAGGCAATGCTCTTTGAAAAACAAATAGGTACTCCCTCGCTCTGAAATGT 1140
 DB 1077 AAGAAAGGCAATGCTCTTTGAAAAACAAATAGGTACTCCCTCGCTCTGAAATGT 1136
 QY 1141 ATACATATGGATTGGACACGGAGACTAAGAAAATGTATAAGTAAATGTAGAGTAAAG 1200
 DB 1137 ATACATATGGATTGGACACGGAGACTAAGAAAATGTATAAGTAAATGTAGAGTAAAG 1196
 QY 1201 AAGAGAAAGAAAGTGGTAAAGTAGCGGACCCACCAATATAATATGATAGATTAG 1260
 DB 1197 AAGAGAAAGAAAGTGGTAAAGTAGCGGACCCACCAATATAATATGATAGATTAG 1256
 QY 1261 AAAAGTAGTTGAAGTAGTGGTGGGATTTTATATAATAAAAAATTTACTATTTTG 1320
 DB 1257 AAAAGTAGTTGAAGTAGTGGTGGGATTTTATATAATAAAAAATTTACTATTTTG 1316
 QY 1321 AGAAGTTTGAATGTATAGAAATTTAGTGGACATCCATAAAGGAAAGTGTATAGAT 1380
 DB 1317 AGAAGTTTGAATGTATAGAAATTTAGTGGACATCCATAAAGGAAAGTGTATAGAT 1376
 QY 1381 TAAATGGGACACAGGAGGTAATACCTTTATGATATATAAATTTTGTATTTTGTATTC 1440
 DB 1377 TAAATGGGACACAGGAGGTAATACCTTTATGATATATAAATTTTGTATTTTGTATTC 1436
 QY 1441 TAAGATTATAATCTATGTTATAATAATAATTTTAAAAATAATACTATATTAATTT 1500

Db 1437 TAAGATTATARAATCTATGTTAATGAATAATAATATTTTAAAAATAAATACTATATTAATT 1496
 Qy 1501 CTGATTAGTCGATTACCGCTTTTATAATTTTCAATACTGAGTAATATGATAAATCAG 1560
 Db 1497 CTGATTAGTCGATTACCGCTTTTATAATTTTCAATACTGAGTAATATGATAAATCAG 1556
 Qy 1561 TTATCTGAAAAGCAATAATATCTTTGTAAACACGCGTTCGGTCAAAATGGGAAGTTCAATG 1620
 Db 1557 TTATCTGAAAAGCAATAATATCTTTGTAAACACGCGTTCGGTCAAAATGGGAAGTTCAATG 1616
 Qy 1621 TGTATTCAATAGTTTTTAATAATAAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA 1680
 Db 1617 TGTATTCAATAGTTTTTAATAATAAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGA 1676
 Qy 1681 AATTTAAATAAATTAATTAGCATGGGAAGTTTCACGGGCATCATTTGACGACACTAGACT 1740
 Db 1677 AATTTAAATAAATTAATTAGCATGGGAAGTTTCACGGGCATCATTTGACGACACTAGACT 1736
 Qy 1741 GTTTCGAACAATGATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1800
 Db 1737 GTTTCGAACAATGATGTCGGGTGTACATCTATGACCTTTCAACTCAAACTAGTGAATAAT 1796
 Qy 1801 GCATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTTCAACG 1860
 Db 1797 GCATCTAGAAATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTTTCAACG 1856
 Qy 1861 GATTGGAATCCTTTTCTAACTTTTAAATAAATAAATAAATGATTAATTTGTAATTTATC 1920
 Db 1857 GATTGGAATCCTTTTCTAACTTTTAAATAAATAAATAAATGATTAATTTGTAATTTATC 1916
 Qy 1921 AACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCAT 1980
 Db 1917 AACACCTCAACATTCATGTTAGCGTACTATAAATAGGTGCTCTTGGTCTCTACTATCAT 1976
 Qy 1981 CACATCAATCTTACACCAACAACTTGAGCTTTAAATTTTCTACTTAATTTCTCAGCAATAAC 2040
 Db 1977 CACATCAATCTTACACCAACAACTTGAGCTTTAAATTTTCTACTTAATTTTCTCAGCAATAC 2036
 Qy 2041 ATTCTAAATATC 2052
 Db 2037 ATTCTAAAGATC 2048

RESULT 6
 AAV15144
 ID AAV15144 standard; DNA; 2042 BP.
 AC AAV15144;
 XX
 XX 02-JUL-1998 (first entry)
 XX
 DE New promoter used for root-specific expression in plants.
 XX Promoter; root; carrot; Kuroda Gosun; root-specific expression;
 KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
 KW 58.
 XX Daucus carota.
 OS
 XX EP824150-A2.
 XX
 XX 18-FEB-1998.
 XX
 XX 12-AUG-1997; 97EP-00113923.
 XX
 XX 12-AUG-1996; 96JP-00212680.
 XX
 XX (SUMO) SUMITOMO CHEM CO LTD.
 XX Torikai S, Oeda K;
 XX
 XX WPI; 1998-122310/12.
 DR
 XX

PT New carrot root gene, promoter and terminator - useful in genetic
 PT engineering for directing root-specific gene expression.
 XX
 XX Claim 2; Page 15-16; 31pp; English.
 XX
 CC The present sequence represents a novel promoter, and is isolated from
 CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
 CC direct root-specific expression in plant cells. Since the promoter
 CC enables expression of a desired protein in the roots of a plant, it is
 CC useful in combat against pathogenic soil fungi and pests which are
 CC difficult to kill by chemicals. It can also be used to improve the
 CC nutritive value of edible root plants
 XX
 SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
 Query Match 11.7%; Score 239.6; DB 2; Length 2042;
 Best Local Similarity 72.6%; Pred. No. 7.7e-35;
 Matches 366; Conservative 0; Mismatches 129; Indels 9; Gaps 4;
 Qy 1554 AATCAGTTATCTGAAAAGCAATAAATATCTTTGTAAACACGCG-----TTCGGTCAAAATG 1609
 Db 1543 AAATTAATTTATCTGAATGATAACATCTTTGTAAACAAAACTGCGCAAAATAGGACCATAA 1602
 Qy 1610 GGAAGTTTCATGTTTCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTATT 1669
 Db 1603 CCAAGTTTCATGTTTCAATAGTTTAAATATAAAGTAAATTTTAAATTAATTTGTTATT 1660
 Qy 1670 TTTCTTTTCAGAAATTTTAAATTAATTTGAGCATGGGAAGTTTCACGGGCATCATTTGAGC 1729
 Db 1661 ATAGTTTAATTTCTTCAATCAATTAACTTTAAATTTGGACATTTATGAGCACTTTATGCC 1720
 Qy 1730 AGCACTAGACTGTTTGAACAATGATGTCGGGTGTACATCTATGACCTTTTCAACTCAAAAC 1789
 Db 1721 CACGTTGTTATTTTAAACAAACGTTTGTCCGGTGTATATTTATGACCTTTTCAACTCAAGC 1780
 Qy 1790 TAGTGAAT-AATGCATTTCTAGATACATCTTTTCAAAATTTTCAAAACAAACACAGCTTTAACT 1848
 Db 1781 TAGCAGTGAATGCTTTCTTGAATATATCTTTTGAATTTTCAAAACAAACAGCACTAACT 1840
 Qy 1849 TTTCTTTTCACGGGATTTGGAATCTTTTCTAAACTTTTAAATTTTAAATTTTAAATTTTAACT 1908
 Db 1841 TTTCTTTTAAACAGATTAGAAATCGTTTCGTAACCTTTTAAATTTTAAATTTTAACT 1898
 Qy 1909 GTAATATTTATCAACACTCAACATTTGATGTTAGCTGCTACTATAAATAGGTGCTCTTGGTG 1968
 Db 1899 ATAATATTTATCAACACTCAACATTTGATGTTAGCTGCTACTATAAATAGGTGCTCTTGGTG 1958
 Qy 1969 CTCTACTATCATCATCAATCTTTTACACCAACCTTTGAGCTTAATTTTCTACTTATT 2028
 Db 1959 CTCTACTATCATCATCAATCTTTTCCAGCAAAACCTTTGAGCTTAATTTTCTACTTATT 2018
 Qy 2029 CTCAGCAATAACATTTCTAAATATC 2052
 Db 2019 TTTAGCAAAAACATTTCTAAAGGTC 2042
 RESULT 7
 AAV15143
 ID AAV15143 standard; DNA; 247 BP.
 AC AAV15143;
 XX
 XX 02-JUL-1998 (first entry)
 XX
 DE New promoter used for root-specific expression in plants.
 KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
 KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
 KW 58.
 XX Daucus carota.
 OS
 XX EP824150-A2.
 PN

XX 18-OCT-2002; 2002JJP-00304115.
XX (SUMO) SUMITOMO CHEM CO LTD.
XX WPI; 2004-360986/34.
XX Expression inducing promoter, useful for expressing foreign gene,
PT comprising first DNA for detecting transcription start point and having
PT minimum promoter function, linked to second DNA having expression
PT inducing promoter function.
XX Disclosure; SEQ ID NO 4; 61pp; Japanese.
XX The invention relates to an expression inducing promoter comprising a
CC first DNA linked to a second DNA at its 5' terminus, where the first DNA
CC has a region which determines a transcription start point of RNA
CC polymerase II and has minimum promoter function, and the second DNA has
CC expression inducing promoter function in a plant cell. The expression
CC inducing promoter is useful for expressing a foreign gene. This sequence
CC represents carrot DNA used in the scope of the invention.
XX Sequence 196 BP; 66 A; 40 C; 19 G; 71 T; 0 U; 0 Other;
SQ
Query Match 9.5%; Score 194.4; DB 12; Length 196;
Best Local Similarity 99.5%; Pred. No. 1.3e-26;
Matches 195; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1853 TTTCACGGATTGGAACTCTTTCTAAACTTTTAAATAAATAAATAAATGCAATTATTGTA 1912
DB 1 TTTCACGGATTGGAACTCTTTCTAACTTTTAAATAAATAAATAAATGCAATTATTGTA 60
QY 1913 TATTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGTGCTCTTGSGTCT 1972
DB 61 TATTATCAACACCTCAACATTGATGTTAGCGTACTATAAATAGTGCTCTTGSGTCT 120
QY 1973 ACTATCATCATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATCTCA 2032
DB 121 ACTATCATCATCAATCTTACACCAACACCTTGAGCTTAATTTTCTACTTATCTCA 180
QY 2033 GCAATAACATTCTAAA 2048
DB 181 GCAATCACAATTCTAAA 196
RESULT 10
ADC56759
ID ADC56759 standard; DNA; 2831 BP.
XX ADC56759;
AC
XX 18-DEC-2003 (first entry)
DT Carrot DNA that encodes an expression inducer type promoter.
DE carrot; expression inducer type promoter; plant; vector; transformant;
KW ds.
KW
XX Daucus carota subsp. sativus.
OS
XX JP2003000252-A.
PN
XX 07-JAN-2003.
PD
XX 31-MAY-2001; 2001JJP-00164069.
PF
XX 31-MAY-2001; 2001JJP-00164069.
PR
XX (SUMO) SUMITOMO CHEM CO LTD.
PA
XX WPI; 2003-472920/45.
DR
XX An expression inducer type promoter derived from DNA of Daucus carota L.

PT var. sativa DC. of 2831 bases and its analogues having the same
PT functions.
XX
XX Claim 1; SEQ ID NO 1; 16pp; Japanese.
XX
XX This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes
CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is the DNA encoding the
CC carrot promoter of the invention.
XX
XX Sequence 2831 BP; 950 A; 456 C; 395 G; 1030 T; 0 U; 0 Other;
SQ
Query Match 7.9%; Score 162.4; DB 10; Length 2831;
Best Local Similarity 82.5%; Pred. No. 1.4e-20;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 22 AGGCCCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTTTGACCTGT 81
DB 1716 AGGCCCTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTTT-ACCGT 1774
QY 82 TTGTATAAGAGTAGAATAATATTTTAAAGCTGCGAATACTTCTCTCACAAC 141
DB 1775 TTGTATAAGAGTAGAATAATATTTTAAAGCTGCGAATACTTCTCTCACAAC 1834
QY 142 TTCGCTCTTTTCCAAACACTTTTATTAACCTTTTCTTCTTCTTCTTCTTCTTCT 201
DB 1835 TTCTGCTTCTTTCCAAACACTTTTATCACTTACCTACTTCTTCTTCTTCTTCT 1894
QY 202 TTGCTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAGATCA 261
DB 1895 TTACTATAAGCAAGAAATCACTTCTTTTAAGCTAACCCAAACGGCCTCAATAAAGATCA 1954
RESULT 11
ADC56761
ID ADC56761 standard; DNA; 2865 BP.
XX ADC56761;
AC
XX 18-DEC-2003 (first entry)
DT Carrot DNA encoding an expression inducer type promoter (SeqID 3).
DE carrot; expression inducer type promoter; plant; vector; transformant;
KW ds.
KW
XX Daucus carota subsp. sativus.
OS
XX JP2003000252-A.
PN
XX 07-JAN-2003.
PD
XX 31-MAY-2001; 2001JJP-00164069.
PF
XX 31-MAY-2001; 2001JJP-00164069.
PR
XX (SUMO) SUMITOMO CHEM CO LTD.
PA
XX WPI; 2003-472920/45.
DR
XX An expression inducer type promoter derived from DNA of Daucus carota L.
PT var. sativa DC. of 2831 bases and its analogues having the same
PT functions.
XX
XX Disclosure; SEQ ID NO 3; 16pp; Japanese.
XX
XX This invention relates to carrot DNA used as an expression inducer type
CC promoter. Specifically, it refers to promoter sequences derived from
CC plant DNA, preferably carrot, and functional mutants thereof that can be
CC used as expression inducers. Furthermore, the present invention describes

CC DNA, vectors, transformants and the process by which to prepare
CC transformants. This polynucleotide sequence is DNA encoding a carrot
CC promoter (SeqID 3) of the invention.

XX SQ Sequence 2865 BP; 963 A; 461 C; 396 G; 1045 T; 0 U; 0 Other;
Query Match 7.9%; Score 162.4; DB 10; Length 2865;
Best Local Similarity 82.5%; Pred. No. 1.4e-20;
Matches 198; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 22 AGGGCCGTGTTGGTTGAGAGAGCAGAGCTGCTTCTGACTTCTTCTTTTGACCTGT 81
DB 1716 AGGGCCGTGTTGTTTATGGAATCAGAGCTGCTTCTGACTTCTTCTTTT-ACCCGT 1774
QY 82 TTGTATAAGAGAGTGAAGATATTTTAAAAAGCTGGAATACCTTCTCTCACAC 141
DB 1775 TTGTGTAAAGCAGAGCAGCTTTTAAAGCTGGAATGCTAGTTTCTCTCACAGC 1834
QY 142 TTCGGCTTCTTTTCCAAACACTTATTAATCTTTTCTTCTTCTTCTTCTTCTTCT 201
DB 1835 TTTCTGCTTCTTTTCCAAACACTTATCACTTACCTTCTTCTTCTTCTTCTTCT 1894
QY 202 TTGCTATAAGCAGAAATCACTTCTTTTAAAGCTAACCCAAAGCGCTCAATAAAGATCA 261
DB 1895 TTACTATAAGCAGAGAGTCAATTTCTTTTAAATTAACCCAAAGCGCCCTTAAGTAATTGA 1954

RESULT 12
AAZ49616/c
ID AAZ49616 standard; DNA; 140 BP.
XX AAZ49616;
XX AC
XX AC
XX AC
XX 07-APR-2000 (first entry)
XX DE Oligonucleotide-4 for synthesis of CR16.1 fragment for plant promoter.
XX KW Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
XX KW soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
XX KW male sterility-related gene; ss.

XX Daucus carota.
XX OS Synthetic.
XX PN EP976832-A2.
XX PD 02-FEB-2000.
XX PF 13-JUL-1999; 99EP-00113732.
XX PR 15-JUL-1998; 98JP-00200372.
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX PI Ishige F, Nishikawa S, Oeda K;
XX WPI; 2000-128374/12.
XX DR Novel promoter used to produce transgenic plants with higher expression
XX PT of a desired gene.
XX PS Disclosure; Page 14; 24pp; English.
XX CC The present sequence is an oligonucleotide (- chain) used to prepare a
XX CC CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX CC fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX CC construction of a plant promoter. The promoter is used for controlling
XX CC the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX CC desaturase and S-locus type specific RNase gene (male sterility-related
XX CC gene) in a host cell especially a microorganism or a plant cell. The
XX CC transformed plant cells can be used to produce transgenic plants. The
XX CC promoter is compact and therefore suitable for higher expression of a
XX CC desired gene in a particular tissue compared to other host tissues

XX SQ Sequence 140 BP; 45 A; 17 C; 34 G; 44 T; 0 U; 0 Other;
Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 9.8e-13;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1918 ATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGTGCTCTTGGTGTCTACTAT 1977
DB 140 ATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGTGCTCTTGGTGTCTACTAT 81
QY 1978 CATCATCAATCTTTACACCAACCTTGAGCTTTAATTTTCTACTTTATTCACGAAT 2037
DB 80 CATCATCAATCTTTCCAGCAACCTTGAGCTTTAATCTTCTACTAATTTTAGCAAA 21
QY 2038 AACATCTTAATATC 2052
DB 20 AACATCTTAAGGTC 6

RESULT 13
AAZ49615
ID AAZ49615 standard; DNA; 140 BP.
XX AAZ49615;
XX AC
XX 07-APR-2000 (first entry)
XX DT
XX DE Oligonucleotide-3 for synthesis of CR16.1 fragment for plant promoter.
XX KW Synthetic DNA; plant promoter; CR16.1 fragment; carrot; transgenic plant;
XX KW soybean glycinin; oligonucleotide; stearyl-ACP-desaturase gene;
XX KW male sterility-related gene; ss.
XX OS Daucus carota.
XX OS Synthetic.
XX PN EP976832-A2.
XX PD 02-FEB-2000.
XX PF 13-JUL-1999; 99EP-00113732.
XX PR 15-JUL-1998; 98JP-00200372.
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX PI Ishige F, Nishikawa S, Oeda K;
XX WPI; 2000-128374/12.
XX DR Novel promoter used to produce transgenic plants with higher expression
XX PT of a desired gene.
XX PS Disclosure; Page 13; 24pp; English.
XX CC The present sequence is an oligonucleotide (+ chain) used to prepare a
XX CC CR16.1 DNA fragment comprising nucleotides 112-246 of the carrot CR16.3
XX CC fragment. The CR16.1 DNA fragment is ligated to a 10 bp synthetic DNA for
XX CC construction of a plant promoter. The promoter is used for controlling
XX CC the expression of a desired gene e.g. soybean glycinin, stearyl-ACP-
XX CC desaturase and S-locus type specific RNase gene (male sterility-related
XX CC gene) in a host cell especially a microorganism or a plant cell. The
XX CC transformed plant cells can be used to produce transgenic plants. The
XX CC promoter is compact and therefore suitable for higher expression of a
XX CC desired gene in a particular tissue compared to other host tissues

XX SQ Sequence 140 BP; 44 A; 34 C; 17 G; 45 T; 0 U; 0 Other;
Query Match 5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. No. 9.8e-13;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
QY 1918 ATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGTGCTCTTGGTGTCTACTAT 1977
DB 140 ATCAACACCTCAACATTTGATGTTAGCTACTATAAATAGTGCTCTTGGTGTCTACTAT 81
QY 1978 CATCATCAATCTTTACACCAACCTTGAGCTTTAATTTTCTACTTTATTCACGAAT 2037
DB 80 CATCATCAATCTTTCCAGCAACCTTGAGCTTTAATCTTCTACTAATTTTAGCAAA 21
QY 2038 AACATCTTAATATC 2052
DB 20 AACATCTTAAGGTC 6

QY 1918 ATCAACACCTCAACATTGATGTTAGCTACTATAATAGGCTCTCTGGTCTCTACTAT 1977
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
5 ATCAACACCTCAACATTGATGTTAGCTACTATAATAGGCTCTCTGGTCTCTACTAT 64
QY 1978 CATCATCAATCTTTACACCAACCTTGAGCTTTAAATTTTCTACTTTCTTCAGCAAT 2037
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
65 CATCATCAATCTTTCCAGCAACACCTTGAGCTTTAAATCTTCTACTAATTTTGTAGCAA 124
QY 2038 AACATCTTAATATC 2052
Db ||||||||||||||||
125 AACATCTTAAGGTC 139

RESULT 14

ID ABZ10246
AC ABZ10246
XX
DT 16-JAN-2003 (first entry)
XX
DE Haematopoietic cell proliferation disorder related DNA sequence #386.
XX
KW Human; haematopoietic cell proliferation disorder; cytostatic;
KW gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia;
KW cytosine methylation state; gene; ds.
XX
OS Homo sapiens.
XX
PN WO200277272-A2.
XX
PD 03-OCT-2002.
XX
PF 26-MAR-2002; 2002WO-EP003401.
XX
PR 26-MAR-2001; 2001US-0278333P.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Berlin K, Braun A, Dietler J, Guetig D, Howe A, Mueller J;
PI Olek A, Piepenbrock C, Adorjan P, Grabs G, Lesche R, Leu E,
PI Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Pelet C;
PI Schwöbe I, Ziebarth H;
XX
DR WPI; 2003-018942/01.
XX
PT Detecting and differentiating between hematopoietic cell proliferative
PT disorders, comprises contacting a target nucleic acid with a reagent that
PT distinguishes between methylated and non-methylated CpG dinucleotides.
XX
PS Claim 28; SEQ ID NO 386; 117pp; English.
XX

CC The present invention describes a method for detecting and
CC differentiating between haematopoietic cell proliferative disorders
CC associated with at least 1 gene and/or their regulatory regions in a
CC subject. The method comprises contacting a target nucleic acid in a
CC biological sample obtained from the subject with at least 1 reagent,
CC which distinguishes between methylated and non-methylated CpG
CC dinucleotides within the target nucleic acid. ABZ09861 to ABZ1118
CC represent specifically claimed nucleotide sequences from the present
CC invention. Oligonucleotides from the present invention can be used; for
CC differentiating between healthy haematopoietic cells and proliferative
CC disorder haematopoietic cells; for differentiating between acute
CC lymphocytic leukaemia and acute myelogenous leukaemia; as probes for
CC determining the cytosine methylation state and/or single nucleotide
CC polymorphisms (SNPs) of haematopoietic cell proliferation disorder
CC related sequences and their complements; and as primers for the
CC amplification of haematopoietic cell proliferation disorder related
CC sequences. The nucleotide sequences from the present invention can also
CC be used for detecting a predisposition to, differentiation between
CC subclasses, diagnosis, prognosis, treatment and/or monitoring of
CC haematopoietic cell proliferative disorders. The present method enables a
CC highly specific classification of haematopoietic cell proliferative

CC disorders allowing for improved and informed treatment of patients
XX
SQ Sequence 8056 BP; 3711 A; 0 C; 371 G; 3974 T; 0 U; 0 Other;
Query Match 4.8%; Score 99.4; DB 8; Length 8056;
Best Local Similarity 44.5%; Pred. No. 6.3e-09;
Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;
QY 548 ATGCTTCTCAAAATGTTTTTATATGTAATAATAATAGCCCAAGGATAGTAAATTC 607
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1975 ATTAATATTAATTTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATTTAAATAT 2034
QY 608 CCGTTTAACCAAGTTTGTATATATATGTTTACACTTACAGAGAGATATTCGTATACTT 667
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2035 ATTTTAAATATTTTATTTTAAATAATTTTATTAATAATTTTAAATATTTATAAT 2094
QY 668 TTAGACGACAGAGACTTAGGTCAAAATGACGCTGGTAAACAGCCCTAGACTTGGTCAC 727
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2095 AATAATTATAAATTTATTAATAATTTTAAATAATTTTATTTTAAATAAATAA 2154
QY 728 TGATAAATAGATAAATTTAGTATAATATAGTAGATCTACAATGACATTTAAATTTAG 787
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2155 TAAAAAATAAATAATTTATTTTAAATAATTTTAAATAATTTTAAATAAATAA 2214
QY 788 CTATTAAATTAAGTTACTAATAATAAGAGAGGTTAGTAAACAGAGCAGGTAAAAACAA 847
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2215 AAAAAAATAAATTTAAATTTTAAATAAATAATTTTATTTTAAATAAATAAATAA 2274
QY 848 GAGCTTGTCTGTGTTTGTAGTTGTTGAGCTCATTTTC-----L-TTAAAAAGTAATGT 901
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2275 AATTTTAAATTTTAAATTTTATTTTAAATAAATTTTAAATTTTGTATTAAATA 2334
QY 902 AAACGTATCTAAACGACATAGAAATTTAGTACAGGTTAAACCTTTTACAGAAATTTATAT 961
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2335 AATTTTAAATTTTAAATAATTTTATTTTATTTTAAATAAATTTTATTTAAATTTTAT 2394
QY 962 TAAACGAAATCATTTTATTAACATGCTCTCGGCTGCATTATTAATAGGATCACTTACT 1021
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2395 TATTTTGTAAATTTTAAATAAATAAATTTTATTTTATTTTAAATAAATAATTT 2451
QY 1022 GATCATCCATTTAAACCTTTGTAAAAACAAATTTCAATGAGATAAATAATCTTTACATGAAA 1081
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2452 TTATTTTATTTATATATTTTATTTTATTTTATTTTAAATAAATAAATAATTTAATGAAA 2511
QY 1082 AGAAGGCAATGCTCTTTTGAATAAACAATAGGTACTCCCTCCGCTCTGAAATGTA 1141
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2512 AATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2571
QY 1142 TACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAAGTAAATGTAGAGTAAAAAGA 1201
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2572 TTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2631
QY 1202 AAGAGAAGAAAAAGTGGTAAAGTAGCGGGACCCCAATAATATAATTTGATAGATTTAGA 1261
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2632 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2691
QY 1262 AAAGTAGTTGAAAGTAGTGGGTGGGTGGATTTTATATATATAAATAATTTACTATTTGA 1321
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2692 TAAATTTTATTTTATTTTATTTTAAATAAATAAATAAATAAATAAATAAATAA 2747
QY 1322 GAAAGTTTGAATAATGTATAGAATTTGAGTGGGACATCCATTAAGGAAAGTGTATAGAATT 1381
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2748 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2807
QY 1382 AAATGGGACAGAGGAGTAATACCTTTATGATATATAAATAATTTTGTATTTTGTATTCAT 1441
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2808 AAATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2865
QY 1442 AAGATTATATCTATGTTATATGATAATATTTTAAATAAATAAATAAATAAATAA 1501
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2866 TTTATGAAAAATAAATTTTGTGTGAAAAATAAATAAATAAATAAATAAATAAATAA 2925
QY 1502 TGATTAGTCGATTACCGCCTTTTATTAATTTTACAATACTGAGTAATGAATAAATCAGT 1561

[illegible]

| | |
|-----------|---|
| RESULT 15 | |
| ABZ10100 | |
| ID | ABZ10100 standard; DNA; 8056 BP. |
| XX | |
| XX | ABZ10100; |
| XX | |
| XX | AC |
| XX | 16-JAN-2003 (first entry) |
| XX | |
| XX | Haematopoietic cell proliferation disorder related DNA sequence #240. |
| XX | |
| XX | Human; haematopoietic cell proliferation disorder; cytostatic; |
| KW | gene therapy; lymphocytic leukaemia; acute myelogenous leukaemia; |
| KW | cytosine methylation state; gene; ds. |
| XX | |
| XX | Homo sapiens. |
| XX | |
| XX | WO200277272-A2. |
| XX | |
| XX | 03-OCT-2002. |
| PD | |
| XX | |
| XX | 26-MAR-2002; 2002WO-EP003401. |
| XX | |
| XX | 26-MAR-2001; 2001US-0278333P. |
| XX | |
| XX | (EPIG-) EPIGENOMICS AG. |
| PA | |
| XX | |
| PI | Berlin K, Braun A, Dietler J, Guetig D, Howe A, Mueller J; |
| PI | Olek A, Pletzenbrock C, Adorjan P, Grabs G, Lesche R, Leu B; |
| PI | Lewin A, Lipscher E, Maier S, Model F, Mueller V, Otto T, Palet C; |
| PI | Schwöpe I, Ziebarth H; |
| XX | |
| XX | WPI; 2003-018942/01. |
| DR | |
| XX | |
| XX | Detecting and differentiating between hematopoietic cell proliferative |
| PT | disorders, comprises contacting a target nucleic acid with a reagent that |
| PT | distinguishes between methylated and non-methylated CpG dinucleotides. |
| XX | |
| XX | Claim 28; SEQ ID NO 240; 117pp; English. |
| PS | |
| XX | |
| XX | The present invention describes a method for detecting and |
| CC | differentiating between haematopoietic cell proliferative disorders |
| CC | |

| | |
|----|---|
| | associated with at least 1 gene and/or their regulatory regions in a |
| | subject. The method comprises contacting a target nucleic acid in a |
| | biological sample obtained from the subject with at least 1 reagent, |
| | which distinguishes between methylated and non-methylated CpG |
| | dinucleotides within the target nucleic acid. AZ09861 to ABZ11118 |
| | represent specifically claimed nucleotide sequences from the present |
| | invention. Oligonucleotides from the present invention can be used: for |
| | differentiating between healthy haematopoietic cells and proliferative |
| | disorder haematopoietic cells; for differentiating between acute |
| | lymphocytic leukaemia and acute myelogenous leukaemia; as probes for |
| | determining the cytosine methylation state and/or single nucleotide |
| | polymorphisms (SNPs) of haematopoietic cell proliferation disorder |
| | related sequences and their complements; and as primers for the |
| | amplification of haematopoietic cell proliferation disorder related DNA |
| | sequences. The nucleotide sequences from the present invention can also |
| | be used for detecting a predisposition to, differentiation between |
| | subclases, diagnosis, prognosis, treatment and/or monitoring of |
| | haematopoietic cell proliferative disorders. The present method enables a |
| | highly specific classification of haematopoietic cell proliferative |
| | disorders allowing for improved and informed treatment of patients |
| XX | |
| SQ | Sequence 8056 BP; 3711 A; 371 C; 371 G; 3603 T; 0 U; 0 Other; |
| | Query Match 4.5%; Score 91.8; DB 8; Length 8056; |
| | Best Local Similarity 44.2%; Pred. No. 1.6e-07; |
| | Matches 627; Conservative 0; Mismatches 767; Indels 25; Gaps 5; |
| QY | 547 AATGCTTCTCAAAATCGTTTTATATGTAAAATAATGCCCATCCAGGAAGAATAAATT 606 |
| DB | 1358 ATTTTATTAATAAAATTAATTTTTTTTTTATAAAAATAATATTCGATTTTTTTTTTATTTT 1417 |
| QY | 607 CCGGTTTAAACGAGTTGGTTAAATATATATGTTTACACTTAACAAGAGATATTCGTAATAC 666 |
| DB | 1418 TTATATTTTTTTTTTTAAAAAATAAAATTAATTTTTTAAATAAATATATTTAA 1477 |
| QY | 667 TTTAGACGACAGAGACTTAGGTCAAATAATGGACGCTGGTAAACAGCCTAGACTTGCTCA 726 |
| DB | 1478 ATAATTTAATTTTAAATTAATATATATAAATAAAAAATAAATACGTTAAAAAATAATTT 1537 |
| QY | 727 CTGATAAATAGATAAATTTGTTTAGTATAATATAGTAGGATCTCACAAATGACATTTAAAT 786 |
| DB | 1538 TTAAATAAATTAATTTTTTTTTTAAAAAATAATATATTTAAATTTTATATTTA 1597 |
| QY | 787 GCATTTAAATTAAGTTACTAATAAATAAGAGAGGTTAGTAAACAGAAGCAGGTA AAAACA 846 |
| DB | 1598 TAATTTTAAATTTTTTAAAAATCGAATAAACGAATCGTAAATTTAAAAAATAATTTT 1657 |
| QY | 847 AGAGCTTGCTGCTGCTGTTTAGTTGAGAGCTCATTTCTTTTAAAGTAATGTGTAAC 906 |
| DB | 1658 AATTAATAC-GATAAAATTTTATTTTTTATTAATAATATAAAAAATTAATTTAAATATAATA 1716 |
| QY | 907 GATCTAAGCACAATAGAAATTTAGTACAGGTTAAAACTTTTACAAGNAATTTATTTAAAC 966 |
| DB | 1717 TATATATTTTAAATTTTAAATATTTTAAAAAATTTAAAAAATTTTAAAAATTTAAAA 1776 |
| QY | 967 GAATAATCATTTTATAACATGCTCTCGCGTGCATTATATAATAGGATCACTTACTGATCA 1026 |
| DB | 1777 AATTAATAAATTAATTAAGTTATATATTTTAAATTAATAAATAAATAATTTTCGTTTA 1836 |
| QY | 1027 TCATTTTAAACCTTGTTTAAAAACAAATTCAAATGAGATAAAATATCTTACAATGAAAAAG 1086 |
| DB | 1837 AAAAAAAAAAAAAAAAAAAAAAATAATTTTAAATTAATTTTATTTTATTTTATTTAT 1896 |
| QY | 1087 GACAAATGCTCTTTGAAAAAACAATAAGGTACTCCCTCCGCTCCCTCTGGAATGTATACAT 1146 |
| DB | 1897 AAAAAATAAAAATTTTATAAAAAATAAATTTAAATTTAATAAATAAATAAATAAT 1956 |
| QY | 1147 ATGGATTTGACACGGAGACTTAAGAAAAATGCTATAAAAGTAATGTAGAGTAAAAAGAGAG 1206 |
| DB | 1957 AATTTTTTAAATAAATAATTAATATATTTTATTTTAAATTTTAAATAAATAATTTTAAAT 2016 |
| QY | 1207 A-----AAGAAAAGTGGGTAAAGTAGCGGGGCCCAATATATATTAATTTGATAGAT 1256 |

Db 2017 ATATTTAAATAAATATATATTTTAAATATATTTTATTTTAAATTTTATTTAAAT 2076
QY 1257 TTAGAAAAGTAGTTGAAAGTAGTGGTGGTGGGATTTTATATATATAAAATTTTACTAT 1316
Db 2077 ATTTTAAATATTTATAAATAAATAATATATAAATTTATTAATATATATTTTAAATTTTAT 2136
QY 1317 TTTGAGAAAGTTTTCGAAATGTATAGAAATGTAGTGGGACATCCATAAAGGAAAGTGATATA 1376
Db 2137 TTTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2196
QY 1377 GAATTTAAATGGACAGAGGAGTAATACCTTTATGATATATAAATTTTGTATTTTGTAT 1436
Db 2197 ACGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2256
QY 1437 TTCATAAGATTAATAATCTATGTTATATAGATAATAATTTTAAATAAATAAATAAATAA 1496
Db 2257 TTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2316
QY 1497 AATTCGTGATTAGTCGATACCGCCTTTTATATTTTACAAATCTGAGTAATAGTAATAA 1556
Db 2317 AATTTTCGTATTTAAATAAATTTTAAATTTTAAATAAATAAATTTTATTTTATATAA 2376
QY 1557 TCAGTTATCTGAAAAGCAAT-----AATATCTTTGTAAACACGCGTTTCGTCAAATGG 1610
Db 2377 TAATTTTAAATTTTATTTTTCGAAAATTTAAATAAATAAATAAATAAATAAATAAATA 2436
QY 1611 GAAGTTCATGTATTCATAGTTTATATATAAAGTAAATTTTAAATTTTAAATTTTAAAT 1666
Db 2437 AAAATTAATTTTATTTTATTTTATTTTATATATATATTTTATTTTATTTTAAATAAAT 2496
QY 1667 ATTTTGTTCAGAAATTTTAAATAAATTTTGAAGCATGGGAGTTTCAGGGCATCATTTG 1726
Db 2497 AAAATTTAAACGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2556
QY 1727 AGCAGCAGTAGACTGTGTTGAACAATGTATGTCGGGTGATCATCTATGACCTTTCAACTCA 1786
Db 2557 AAAATTTAAATAAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2616
QY 1787 AACTAGTGAATAATGCATCTTAGAATAACATCTTTTCAAAATTTCAACAAACACAGCTTTAA 1846
Db 2617 AATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2676
QY 1847 CTTTTCTTTCAACGGATTCGATCCTTTTCTAAACCTTTTAAATAAATAAATAAATAA 1906
Db 2677 AT-----TTTTAAATTTTAAATTTTATTTTATTTTATTTTATTTTAAATAAATAA 2732
QY 1907 TTGTAATATTTTCAACACCTCAACATTTGATGTTAGCGT 1945
Db 2733 AATAAATTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2771

Search completed: December 6, 2004, 19:02:26
Job time : 1138.79 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:08:52 ; Search time 205.69 Seconds
(without alignments)
7090.965 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtgcctacagcaca.....gcaataacattcttaaatc 2052

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/prodata/1/ina/5A COMB.seq.*

2: /cgn2_6/prodata/1/ina/5B COMB.seq.*

3: /cgn2_6/prodata/1/ina/6A COMB.seq.*

4: /cgn2_6/prodata/1/ina/6B COMB.seq.*

5: /cgn2_6/prodata/1/ina/PCBUS COMB.seq.*

6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------------|
| 1 | 241.2 | 11.8 | 2042 | 2 | US-08-911-434A-2 |
| 2 | 196.6 | 9.6 | 247 | 2 | US-08-911-434A-1 |
| 3 | 195.6 | 9.5 | 246 | 3 | US-09-352-608-2 |
| 4 | 119 | 5.8 | 140 | 3 | US-09-352-608-6 |
| 5 | 119 | 5.8 | 140 | 3 | US-09-352-608-7 |
| 6 | 69.4 | 3.4 | 7218 | 1 | US-08-232-463-14 |
| 7 | 68.6 | 3.3 | 1141 | 4 | US-09-806-708B-22 |
| 8 | 68.2 | 3.3 | 1141 | 4 | US-09-806-708B-22 |
| 9 | 61.6 | 3.0 | 20674 | 4 | US-09-641-638-651 |
| 10 | 61.6 | 3.0 | 20674 | 4 | US-10-170-097-651 |
| 11 | 60.2 | 2.9 | 19124 | 2 | US-08-487-826B-13 |
| 12 | 59.6 | 2.9 | 7664 | 4 | US-10-204-708-84 |
| 13 | 59.6 | 2.9 | 53332 | 4 | US-09-801-861-3 |
| 14 | 59.6 | 2.9 | 53332 | 4 | US-10-224-562-3 |
| 15 | 58.6 | 2.9 | 5152 | 4 | US-10-204-708-73 |
| 16 | 58.2 | 2.8 | 5852 | 1 | US-07-867-106-2 |
| 17 | 58 | 2.8 | 8920 | 2 | US-08-446-855A-1 |
| 18 | 58 | 2.8 | 8920 | 3 | US-09-150-741-1 |
| 19 | 58 | 2.8 | 10144 | 4 | US-10-204-708-94 |
| 20 | 57 | 2.8 | 640681 | 4 | US-09-790-988-1 |
| 21 | 56.8 | 2.8 | 6583 | 4 | US-10-204-708-25 |
| 22 | 56.4 | 2.7 | 20674 | 4 | US-09-641-638-651 |
| 23 | 56.4 | 2.7 | 20674 | 4 | US-10-170-097-651 |
| 24 | 56 | 2.7 | 6113 | 4 | US-10-204-708-13 |
| 25 | 55.8 | 2.7 | 658 | 3 | US-08-998-416-595 |
| 26 | 55.8 | 2.7 | 6306 | 4 | US-10-204-708-49 |
| 27 | 55.4 | 2.7 | 19124 | 2 | US-08-487-826B-13 |

| | | | | | | |
|----|------|-----|--------|---|--------------------|-------------------|
| 28 | 54.6 | 2.7 | 640681 | 4 | US-09-790-988-1 | Sequence 1, Appli |
| 29 | 54.2 | 2.6 | 1055 | 4 | US-09-806-708B-23 | Sequence 23, Appl |
| 30 | 54.2 | 2.6 | 6317 | 4 | US-10-204-708-11 | Sequence 11, Appl |
| 31 | 53.8 | 2.6 | 1055 | 4 | US-09-806-708B-23 | Sequence 23, Appl |
| 32 | 53.4 | 2.6 | 6152 | 3 | US-08-973-462-1 | Sequence 1, Appli |
| 33 | 53.2 | 2.6 | 6156 | 4 | US-10-204-708-59 | Sequence 59, Appl |
| 34 | 53.2 | 2.6 | 168575 | 4 | US-09-426-290-1 | Sequence 1, Appli |
| 35 | 53 | 2.6 | 636 | 3 | US-08-998-416-1137 | Sequence 1137, Ap |
| 36 | 53 | 2.6 | 53332 | 4 | US-09-801-861-3 | Sequence 3, Appli |
| 37 | 53 | 2.6 | 53332 | 4 | US-10-224-562-3 | Sequence 3, Appli |
| 38 | 52.8 | 2.6 | 782 | 3 | US-08-998-416-224 | Sequence 224, App |
| 39 | 52.8 | 2.6 | 6243 | 2 | US-09-056-075-1 | Sequence 1, Appli |
| 40 | 52.8 | 2.6 | 319608 | 4 | US-09-539-333D-1 | Sequence 1, Appli |
| 41 | 52.8 | 2.6 | 319608 | 4 | US-09-679-409-1 | Sequence 1, Appli |
| 42 | 52.6 | 2.6 | 8920 | 2 | US-08-446-855A-1 | Sequence 1, Appli |
| 43 | 52.6 | 2.6 | 8920 | 3 | US-09-150-741-1 | Sequence 1, Appli |
| 44 | 52.4 | 2.6 | 636 | 3 | US-08-998-416-1137 | Sequence 1137, Ap |
| 45 | 52.4 | 2.6 | 1500 | 4 | US-09-601-198-36 | Sequence 36, Appl |

ALIGNMENTS

RESULT 1
US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Daucus carota L.
; INDIVIDUAL ISOLATE: Kuroda Gosun
; FEATURE:
; NAME/KEY: promoter
; LOCATION: 1..2042
; US-08-911-434A-2

Query Match 11.8%; Score 241.2; DB 2; Length 2042;
Best Local Similarity 72.8%; Pred. No. 7.3e-42;

| Matches | 367; | Conservative | 0; | Mismatches | 128; | Indels | 9; | Gaps | 4; |
|---------|------|---|------|------------|------|--------|----|------|----|
| Qy | 1554 | AAATCAGTATCTGAAAAGCAATAATATCTTTGTTAAACAGCG- - - - -TTCCGGTCAATG | 1609 | | | | | | |
| Db | 1543 | AAATATTTATCTGATGATTAACATCTTTGTTAAACAATACTGGCCAAATAGGACCAATAA | 1602 | | | | | | |
| Qy | 1610 | GGAGGTTTCATGTGTATTCAATAGTGTTTTAAATATAAAAGCTAAATTTTAAATTTGTTTATT | 1669 | | | | | | |
| Db | 1603 | CCAAGTTCACGTGTATTCTAAAAATTAACTAACTAAACATGAGTATTTTCTTT - -TTCAAGGT | 1660 | | | | | | |
| Qy | 1670 | TTTGTGTTCAAGAAATTTAAAAATAATTAATTGAGCATGGAAGTTTCACGGGCAATCATTTGAGC | 1729 | | | | | | |
| Db | 1661 | ATAAGTTAAATCTTCAATCAATTAACATTTAAATTTGGACATTAATTGAGCAACCTTTATGCC | 1720 | | | | | | |
| Qy | 1730 | AGCACTAGACGTGTTTGAACAATGATGTCGCGGTGTACATCTATGACCTTTCAACTCAAC | 1789 | | | | | | |
| Db | 1721 | CAGTTGTAATGTTTAAACAACGTTTGTGCGGTGTATATTATGACCTTTCAACTCAAGC | 1780 | | | | | | |
| Qy | 1790 | TAGTGAAT - AATGCATTTCTAGAATACATCTTTTCAAAATTTCAACAACACAGACTTTAACT | 1848 | | | | | | |
| Db | 1781 | TAGCCAGTGAATGCTTTCTAGATATATCTTTTGAATTTTCAACAACACAGACACTAACT | 1840 | | | | | | |
| Qy | 1849 | TTTCTTTCAACGGGATTGGAAATCCTTTTCTAAACCTTTTTTAAAAATAAAAAATGCATTAAT | 1908 | | | | | | |
| Db | 1841 | TTTCTTTTAAACAGATTAGAATCGTTTCTTAAACCTTTTAAAAAT - -AAAAAATACATTAAT | 1898 | | | | | | |
| Qy | 1909 | GTAATATTTATCAACACCTCAACATTCATGTTAGCGTACTATAATATAGGTGCTCTGGTG | 1968 | | | | | | |
| Db | 1899 | ATAATATTTATCAACACCTCAACATTCATGTTAGCGTACTATAATATAGGTGCTCTGGTG | 1958 | | | | | | |
| Qy | 1969 | CTCTACTATCATCACATCAATCTTTACACCAACACCTTTGAGCTTAAATTTTTTCTACTTTATT | 2028 | | | | | | |
| Db | 1959 | CTCTACTATCATCACATCAATCTTCCAGCAACAACTTTGAGCTTAATCTTTTCTACTAATTT | 2018 | | | | | | |
| Qy | 2029 | CTCAGCAATTAACATTTCTAAATATC | 2052 | | | | | | |
| Db | 2019 | TTTAGCAAAAAACATTTCTAAAGGTC | 2042 | | | | | | |

RESULT 2

```

US-08-911-434A-1
; Sequence 1, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,434A
; FILING DATE: 12-AUG-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Raymond C.
; REGISTRATION NUMBER: 21,066
; REFERENCE/DOCKET NUMBER: 2185-0199P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)205-8000
; TELEFAX: (703)205-8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

[illegible]

DECLASS. AUTHORITY: 25

```

RESULI_3
US-09-352-608-2
; Sequence 2, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Daucus carota
US-09-352-608-2

Query Match          9.5%; Score 195.6; DB 3; Length 246;
Best Local Similarity 89.5%; Pred. No. 1.8e-32;
Matches 222; Conservative 0; Mismatches 24; Indels 2; Gaps 1;

Qy      1805  TCTAGAATACACTCTTTTCAAAATTTCAACAAACACACGCTTTAACTTTTCTTTCAACGGAATT 1864
Db      1      TCTAGAATATATCTTTTGAAATTTCAACAAACACACGACTAACTTTTCTTTTAAACAGATT 60

Qy      1865  GGAATCCTTTTCTAAACTTTTTTAAAAATAAAAAATGCAATTATGTAAATATTTATCAACA 1924
Db      61     AGAATCGTTTCTCTAAACTTTTAAAAAT--AAAAAATACATTTACTATAATATTTATCAACA 118

```

| | | | |
|----|------|--|------|
| Qy | 1925 | CCTCAACATTGATGGTTAGCGTCTATATAAATAGGTCTCTTTGGTGCTCTACTATCATCATCA | 1984 |
| Db | 119 | CCTCAACRTTCATGTT | 178 |
| Qy | 1985 | TCAATCTTACACCAAAACCTTGAGCTTAATTTTTTCTACTATTCTCAGCAATAACAATTC | 2044 |
| Db | 179 | TCAATCTTCAGACACAAACCTTGAGCTTAATCTTTCTACTTAATTTTTTAGCAAAACATTC | 238 |
| Qy | 2045 | TAAATATC | 2052 |
| Db | 239 | TAAAGGTC | 246 |

```

RESULT 4
US-09-352-608-6
; Sequence 6, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2185-0353P
; CURRENT APPLICATION NUMBER: US/09/352,608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 140
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-6

```

| | Query Match | 5.8%; | Score 119; | DB 3; | Length 140; |
|----|-----------------------|--|--------------------|-----------|-------------|
| | Best Local Similarity | 92.6%; | Pred. No. 2.1e-16; | | |
| | Matches 125; | Conservative 0; | Mismatches 10; | Indels 0; | Gaps 0; |
| Qy | 1918 | ATCAACACCTCCACATTGATGTTAGCGTACTATAAATAGGTGCCTCTGGTCTCTACTAT | 1977 | | |
| | | | | | |
| Db | 5 | ATCAACACCTCCAACTTCATGTTAGCGTACTATAAATAGGTGCCTCTGGTCTCTACTAT | 64 | | |
| | | | | | |
| Qy | 1978 | CATCACATCAATCTTTACACACAAACCTTCAGCTTAATTTTTTCTATTATTCAGCAAT | 2037 | | |
| | | | | | |
| Db | 65 | CATCACATCAATCTTTCAGCACAAACCTTCAGCTTAATCTTCTACTAAATTTTTCAGCAA | 124 | | |
| | | | | | |
| Qy | 2038 | AACATTCTAAATATC | 2052 | | |
| | | | | | |
| Db | 125 | AACATTCTAAAGGTC | 139 | | |
| | | | | | |

```

RESULT 5
US-09-352-608-7/c
; Sequence 7, Application US/09352608
; Patent No. 6218598
; GENERAL INFORMATION:
; APPLICANT: ISHIGE, Fumiharu
; APPLICANT: NISHIKAWA, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: Plant Promoter
; FILE REFERENCE: 2195-0353P
; CURRENT APPLICATION NUMBER: US/09/352, 608
; CURRENT FILING DATE: 1999-07-13
; EARLIER APPLICATION NUMBER: 10-200372 JAPAN
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 140
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-352-608-7

Query Match      5.8%; Score 119; DB 3; Length 140;
Best Local Similarity 92.6%; Pred. NO. 2.1e-16;
Matches 125; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1918 ATCAACACCTCAACATTGATGTTAGCGGTACTATAAATAGTGTCTTTGGTGCTCTACTAT 1977
Db 140 ATCAACACCTCAACATTCATGTTAGCGGTACTATAATAGTGTCTTTGGTGCTCTACTAT 81

Qy 1978 CATCACATCAATCTTACACCAACACCTTGAGCTTAATTTTTTCTACTTATTCACAGCAAT 2037
Db 80 CATCACATCAATCTTCCAGCACAAACCTTGAGCTTAATCTTCTACTAATTTTTCAGCAA 21

Qy 2038 AACATTCTAAATATC 2052
Db 20 AACATTCTAAAGTC 6

```

RESULT 6
 US-08-232-463-14/c
 ; Sequence 14, Application US/08232463
 ; Patent No. 5670367
 ; GENERAL INFORMATION:
 ; APPLICANT: DORNER, P.
 ; APPLICANT: SCHEIFLINGER, F.
 ; APPLICANT: FALKNER, F. G.
 ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
 ; NUMBER OF SEQUENCES: 52
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Foley & Hardner
 ; STREET: 1800 Diagonal Road, Suite 500
 ; CITY: Alexandria
 ; STATE: VA
 ; COUNTRY: USA
 ; ZIP: 22313-0299
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/232,463
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION NUMBER: US/07/935,313
 ; FILING DATE:
 ; APPLICATION NUMBER: EP 91 114 300.6
 ; FILING DATE: 26-AUG-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: BENT, Stephen A.
 ; REGISTRATION NUMBER: 29,768
 ; REFERENCE/POCKET NUMBER: 30472/114 IMMU
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (703)836-9300
 ; TELEFAX: (703)683-4109
 ; TELEX: 899149
 ; INFORMATION FOR SEQ ID NO: 14:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 7218 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; IMMEDIATE SOURCE:
 ; CLONE: pTZgpt-F1s
 US-08-232-463-14

Matches 16; Conservative 212; Mismatches 123; Indels 0; Gaps 0;
QY 1051 ATTCAATGAGATAAATATCTTTCAATGAAAGAGACAATGTCTCTTTGAAAAACAA 1110
D 1461 AGTAAAGAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRR 1402
QY 1111 ATAGGTACTCCCTCGCTCCCTGCTGAAATGATATACATATGATGAGTACAGGAGATAAGA 1170
D 1401 RRR 1342
QY 1171 AAAATGATATAAGTATAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1230
D 1341 RRR 1282
QY 1231 GACCCACCAATATAAATTTAGATAGATTTAGAAAGTAGTTGAAAGTAGTGGTGGTGG 1290
D 1281 RRR 1222
QY 1291 ATTTTATATATAAATTTACTATTTTGAGAAAGTTTGAATGATATAGATTTAGTG 1350
D 1221 RRR 1162
QY 1351 GCACATCCATAAAGAGAGTCTATAGAAATTAATGAGGACAGGAGAGTAA 1401
D 1161 RRR 1111

RESULT 7

US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAE1 promoters
US-09-806-708B-22

Query Match 3.3%; Score 68.6; DB 4; Length 1141;
Best Local Similarity 13.1%; Pred. No. 1.3e-05;
Matches 141; Conservative 392; Mismatches 536; Indels 10; Gaps 5;
QY 538 TTTTGGCGAATGCTTCTCAAAATGTTTTTATATGTAATAATGCGCATCCCAAGGATA 597
D 1129 TTTTCTTKYKANNNNNNNGMGKMDATKWSATGTAWTTHAKRGATMCWYWTGTGN 1070
QY 598 AGTAAATTCCTGTTTAAACAGTTGTTAATATATATGTTTACACTTCAAGAGATAT 657
D 1069 RWCWRTYMRWYTRSNWNSCATKBMWTKWTATKTRTH-WYAMCANNNNNNWCAT 1011
QY 658 CTAATACTTTTACAGCAAGAGACTTAGGTCAAAAATGGACGCTGGTGAACAGCCTAG 717
D 1010 NGYAKSCATNNAWYATTRWAAYAAAKWARWAGNNRMVYGAAGNKGWCAAAWATMGBW 951
QY 718 ACTTGCTACTGATAAATAGATAAATGTTAGTATATATAGTAGGATCTACATGACATT 777
D 950 ADTAGKMNNNNNNTTDDRMM--KAKNNNNNNAWYATCYNRAATNNKATTHMKWTH 893
QY 778 AAAATTAGAGCTATTAATTAAGTACTATAATAAGAGAGGTTAGTAAACAGAGAGAG 837

Db 892 GAHSKRTRHHTRCRRTKYNNNNNNNNNARTVYVYHHAARRWNAWTRTNNNNNNNNNAC 833
QY 838 GTAAAAACAAGAGCTGCTGCTGTGTTAGTTGTTGTGAGCTCATTTCTTTTAAAAAGTA 897
D 832 RNRTRW--WABWKHSWCNN 775
QY 898 ATGTAATCTGATCTAAAGCAGATAGAAATTTAGTACAGGTTTAAACACTTTTACAAGAA 957
D 774 NNYMHAAVTTTHTDWCYKTMWNTWYDMMTMTBTTTNNMTTSTNNNNNNNNNNNNNN 715
QY 958 ATATTAAACGAAATCATTTTATAACATGCTCTCGGCTGCTCATTTATAATAGGATCACT 1017
D 714 NNNNNWKAAYAHATNNNGCWNNDARTNNNTTVRRRWMTNTKTRWSTTR--HHYT 658
QY 1018 TACTGATCATCCATTAACAACTTGTAAACCAATTCATAGATAGATAAATATCTTCAAT 1077
D 657 GATNNNNNNNNNNNNNNNSCCTCTRMWTRWTKMGDMTVRKVKWDRDCTCTYDVWAD 598
QY 1078 GAAAAGAGGACAATGCTCTTTGAAAAACAAATAGTACTCCCTCGCTCCCTCTGAAA 1137
D 597 SWVWYANMRECRDVTYTRNNYCKSYAHSYWYNNAMWYRRYSARNWSSMARWTTTRNN 538
QY 1138 TGTATACATATGGATTGGACACGGAGACTAAGAAAAATGTATAAGTAATGTAGAGTAAA 1197
D 537 WMSGBVWRWAGTWWRHNNNNNTDTRYWWKWARBITTVYVDSMCNAKSMWRGNWR 478
QY 1198 AAGAAAGAGAGAAAGTGGGTAAAGTACGGGACCCCAACCAATATATAATTTGATAGATT 1257
D 477 AMKWWAANNDAGAMDHWTYWNGNNTWMRRRAWKMMNAWCCRAYCCNNNNNNRACVHKH 418
QY 1258 TAGAAAAGTAGTGAAGTAGTGGTGGGAGTTTATATATATAAAAAATTTACTATT 1317
D 417 KMRWRTWKYM--WKAACNNNNNNBKAMYRMVAMMYSRDNTTNDMMWTSDBWHMYTVDTY 360
QY 1318 TTGAGAAAGTTTTCGAAATGTATAGAAATTTGAGTGGGACATCCATAAAGAGAGTATAG 1377
D 359 MRAWNNNNNNNNWRECKTTSWMMWMDHNNHCTTGNWGSAYBMAAAMSWAAGASNBVT 300
QY 1378 AATTAATAGGACAGAGAGGAGTAACTCTTTATGATATATAAAATTTTGTATTGTTGATT 1437
D 299 YNWCWRTYMGKTMTNNNNNNKAWYRTKTVAMCNRYYDTAVMTBKRNYKYCYAYBWY 240
QY 1438 TCATAAGATTAATAATCTATGTTAATGATAATAATTTTAAAAATAACTATATA 1497
D 239 YBMYGKHWHBWRBRHSNNWVCRNKYMYVSHYHAMRYBKWABAVGCNNNNNNWKDM 180
QY 1498 ATCTGATTAGTACGCTACCGCTTTTATAATTTTACAATACTAGTAAATATGATAAAT 1557
D 179 AHHHCATNNNNNNNNWYAYMHMHKKGKAAWNNKTBABDDHBAHVKYTWYRYYWCM 120
QY 1558 CAGTTATCTGAAAAGCAAAATAATATCTTTGTAACACAGCGTTCGGTCAAAATGGGAAGTT 1616
D 119 CWNNAKAVRTAMKHMWYTDYVVSANNTGVRWMMWRWCMWYNNNNNNNNNNNNNNNN 61

RESULT 8

US-09-806-708B-22
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence

```
;
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PaeI promoters
US-09-806-708B-22

Query Match      3.3%; Score 68.2; DB 4; Length 1141;
Best Local Similarity 12.6%; Pred. No. 1.6e-05;
Matches 137; Conservative 405; Mismatches 531; Indels 13; Gaps 5;

QY 627 ATATATATGTTTACACTTACAAGAGGATATTCGTAATACCTTTTAGAGCACAAGACTTA 686
DB 57 RKWTARMYCKYRRYNNKSRWGWYKKYWCANNTSBRVHARRKDKMTAYBMTN 116
QY 687 GGTCAAAAATGGACGCTGTAACAGCCTGCTGCTCACTCATATAAGATAATGTT 746
DB 117 KWGKTGRRHYWRWABDVTVDHYVTAMNNATWTCMDKDDKPRRWKKNNAATGWD 176
QY 747 AGTATAATATAGTAGGACTTACAATGACATATAAATTAGAGCTATTAATTAAGTTACTAA 806
DB 177 DDTKYHMNNNGCBTVMVRYKTRDWSBKRMNYGMBWKNWSYDVTVYVWVWDDMKCR 236
QY 807 TAAATAAGAGAGTTAGTAAACAGAAAGCGGTAAACACAGAGCTTGTCTGTGTGTT 866
DB 237 KVRWRVTRGRMRNYMV-AMBTARRRYNNGWTBMAVYRRWTNNNNNNAKAMCKRAKYW 295
QY 867 TAGTGTGTGAGCTCATTTCTTTTAAAGAGTAATGTAACCTGATCTAAAGCACATAGAAAT 926
DB 296 GWNRAVNSTCTTWKSKTKYRTSVANNCRAGDANKHKWKSAMGYVNNNNNNNW 355
QY 927 TTAGTACAGTTAAAACTTTTACAAGAATTTATTAACAGCAAAATCATTTTATAACATG 986
DB 356 TYKKARHBARWDVWHSASWKWHANAHAHYSRKKWBYKRTVMVNNNGTTTMMKRWMAWYW 415
QY 987 TCTCTCGGCTGTC----ATATAATAGGATACCTTACTGATCATCATTAACCTGT 1042
DB 416 KMDMDWBGTYNNNNNGRTYGTGKNQWYTYKWKANNCKWRADHKTCTHNTTMMWK 475
QY 1043 TAAACAAATTCATGAGATAAATATCTTACAAGAAAGAGGACAATGCTCTTTGCA 1102
DB 476 KTYWNNCYKWSMTNGSHRBAALAVTYWVWVRYAHANNNDWYVWKACTWYKYBVCSK 535
QY 1103 AAAAAAATAGTACTCCCTCGCTCCCTCTGAAATGTATACATATGGAATGGACACGGA 1162
DB 536 WNNYAAWYTKSSWNTSYRYKTNNSWRSDTRSMGRANNYARABHYGKW-NTRWW 594
QY 1163 GACTAAGAAAAATGTATAAGTAATGTAGATGTAAGAAAGAGAAAGAAAGAGTGGTAA 1222
DB 595 BWSHTWBHERAGAAHYWMBMYBAKCHCMKAWYKAKKYAGAGGSGNNNNNNNNNNNN 654
QY 1223 AGTAGCGGACCCACCAATATATAATTGATAGATTTAGAAAAAGTAGTTCGAAAGTAGTGG 1282
DB 655 ATCARDYVAAASHWYANAKWYKYBAANNAYTHANNWGCWNNATDTRTMMKNNNN 714
QY 1283 TGGGTGGGATTTTATATATAAAAAATTTACTATTTTGAAGAAAGTTTGAAGATGTATAGA 1342
DB 715 NNAAGTWKNNNNNNAKASAAKNYAAAANKHGWKANKWAMRGWHAADAAABTTDKR-- 772
QY 1343 ATTGAGTGGGACATCCATAAAGGAAAGTGTATAGAAATTAATCGGACAGAGGAGTAAT 1402
DB 773 -NNGAYTKYTTTTNNNNYTRGVVNTAARDGWANNNNNNNNNNNNNNNNNNNNNNNN 831
QY 1403 ACCTTTATGATATATAAATTTTGTATTTTGTGATTTTCAAGATTAATAATCATGTTAT 1462
DB 832 YGTNNNNNNNNNAYAWTKWYTTTDDRRBRBATNNNNNNNNNNNNNNNNNNNNNNNN 891
QY 1463 AATGATAATATAATTTTAAAAATAATACATATATAATTTCTGATTCGATTCACGCCCTT 1522
DB 892 CDAWKKWDATKMNNA TTYNRGTAWRTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 951
QY 1523 TTATAATTTTCAATACTAGTAAATGATGATGATGATGATGATGATGATGATGATGAT 1578
DB 952 WVCATKTKTKGCMNCTTTCRKYKNNCTWYTWTTTTRTTTAAATRWKTNNTGSMTRCNA 1011
```

```
QY 1579 ATATCTTTGTTAAAAACAGCGTTCGGTCAAAATGGAACTTCATGTGTTTCAATAGTTTAA 1638
DB 1012 TGWKNNNYVWGTWRTATYATMATRWKAWKVMATGSMWNTNSYARWAYTKRAYKGYWA 1071
QY 1639 TATAAAGTAAATTTTAAATTAATTTGTTTATTTTGTTCAGAAAATTTAAAAATAAATTT 1698
DB 1072 CAWRWRGKATCYMTDNAMWTACATSWATHKYNNHMKCNKNNNNNNNNNNNNNNNN 1131
QY 1699 GAGCAT 1704
DB 1132 CDGARY 1137

RESULT 9
US-09-641-638-651
; Sequence 651, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CPI
; CURRENT APPLICATION NUMBER: US/09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 651
; LENGTH: 20674
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1123..3123
; OTHER INFORMATION: 5'regulatory region
; NAME/KEY: exon
; LOCATION: 3124..3297
; OTHER INFORMATION: exon 1
; NAME/KEY: exon
; LOCATION: 3871..4072
; OTHER INFORMATION: exon 2
; NAME/KEY: exon
; LOCATION: 5552..5633
; OTHER INFORMATION: exon 3
; NAME/KEY: exon
; LOCATION: 5758..5880
; OTHER INFORMATION: exon 4
; NAME/KEY: exon
; LOCATION: 5996..6099
; OTHER INFORMATION: exon 5
; NAME/KEY: exon
; LOCATION: 6349..6509
; OTHER INFORMATION: exon 6
; NAME/KEY: exon
; LOCATION: 7379..7522
; OTHER INFORMATION: exon 7
; NAME/KEY: exon
; LOCATION: 8645..8854
; OTHER INFORMATION: exon 8
; NAME/KEY: exon
; LOCATION: 12254..12340
; OTHER INFORMATION: exon 9
```

```
NAME/KEY: exon
LOCATION: 12854..13023
OTHER INFORMATION: exon 10
NAME/KEY: exon
LOCATION: 13308..13429
OTHER INFORMATION: exon 11
NAME/KEY: exon
LOCATION: 16567..16667
OTHER INFORMATION: exon 12
NAME/KEY: exon
LOCATION: 16775..16945
OTHER INFORMATION: exon 13
NAME/KEY: exon
LOCATION: 17063..17554
OTHER INFORMATION: exon 14
NAME/KEY: misc.feature
LOCATION: 17555..20674
OTHER INFORMATION: 3'regulatory region
NAME/KEY: allele
LOCATION: 1128
OTHER INFORMATION: 10-508-191 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1182
OTHER INFORMATION: 10-508-245 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1559
OTHER INFORMATION: 10-509-284 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 1570
OTHER INFORMATION: 10-509-295 : deletion of C
NAME/KEY: allele
LOCATION: 1827
OTHER INFORMATION: 10-510-173 : variable motif ATTGA or TTTTTT
NAME/KEY: allele
LOCATION: 2048
OTHER INFORMATION: 10-511-62 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2323
OTHER INFORMATION: 10-511-337 : insertion of T
NAME/KEY: allele
LOCATION: 2341
OTHER INFORMATION: 10-512-36 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 2623
OTHER INFORMATION: 10-512-318 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2832
OTHER INFORMATION: 10-513-250 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2844
OTHER INFORMATION: 10-513-262 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 2934
OTHER INFORMATION: 10-513-352 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 2947
OTHER INFORMATION: 10-513-365 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 3802
OTHER INFORMATION: 12-206-81 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 4062
OTHER INFORMATION: 10-343-231 : deletion of C
NAME/KEY: allele
LOCATION: 4088
OTHER INFORMATION: 12-206-366 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4109
OTHER INFORMATION: 10-343-278 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 4170
OTHER INFORMATION: 10-343-339 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 5903
OTHER INFORMATION: 10-346-23 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6019
OTHER INFORMATION: 10-346-141 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6141
OTHER INFORMATION: 10-346-263 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6183
OTHER INFORMATION: 10-346-305 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6338
OTHER INFORMATION: 10-347-74 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6375
OTHER INFORMATION: 10-347-111 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 6429
OTHER INFORMATION: 10-347-165 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 6467
OTHER INFORMATION: 10-347-203 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6484
OTHER INFORMATION: 10-347-220 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 6534
OTHER INFORMATION: 10-347-271 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 6611
OTHER INFORMATION: 10-347-348 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 7668
OTHER INFORMATION: 10-348-391 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8608
OTHER INFORMATION: 10-349-47 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 8658
OTHER INFORMATION: 10-349-97 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 8703
OTHER INFORMATION: 10-349-142 : polymorphic base G or C
NAME/KEY: allele
LOCATION: 8777
OTHER INFORMATION: 10-349-216 : deletion of CTG
NAME/KEY: allele
LOCATION: 8785
OTHER INFORMATION: 10-349-224 : polymorphic base G or T
NAME/KEY: allele
LOCATION: 8926
OTHER INFORMATION: 10-349-368 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12171
OTHER INFORMATION: 10-350-72 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 12429
OTHER INFORMATION: 10-350-332 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13341
OTHER INFORMATION: 10-507-170 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 13492
OTHER INFORMATION: 10-507-321 : polymorphic base A or C
NAME/KEY: allele
LOCATION: 13524
OTHER INFORMATION: 10-507-353 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 13535
```

Query Match

3.0%; Score 61.6; DB 4; Length 20674;
Best Local Similarity 48.3%; Pred. No. 0.00075;


```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israeleen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

```

| Query Match | 2.9%; | Score 60.2; | DB 2; | Length 19124; |
|-----------------------|-----------------|--|-----------|---------------|
| Best Local Similarity | 49.2%; | Pred. No. 0.0014; | | |
| Matches 271; | Conservative 0; | Mismatches 273; | Indels 7; | Gaps 4; |
| Qy | 1162 | AGACTAAGAAAAATGTATAAAGTAATCTAGAGTAAAGAAAGAGAAAGAAAGTGGGTA | 1221 | |
| Db | 1438 | ATAATTATGAHAAAAATATATTTAATATTTTATTTAATTTTAAAAAAGAGAGAGA | 1379 | |
| Qy | 1222 | AAGTAGCGGACCCACCAATATATAATTTGATAGATTTTAGAAAAAGTAGTTGAAAGTAGTGG | 1281 | |
| Db | 1378 | ACAATTAATAA --- AAAAAGATATTAATATATATACTATTAGATACCGATTGTTAAGT | 1322 | |
| Qy | 1282 | GTGGTGGGATTTTATATTTATATAAAAAATTTACTATTTTGTGAAAAAGTTTTGAAATGTATAG | 1341 | |
| Db | 1321 | TTTCTTTAGTCTTGTATGTTATACAAATATA - TTTTATTAATAATTTTACATATTTAATA | 1263 | |
| Qy | 1342 | AAATTGAGTGGGACATCCATAAAAAGAAAGTGTAAGAAATTTAAATGGGACAGAGGGAGTAA | 1401 | |
| Db | 1262 | TTTTTGATAACATATACCTAAATATATATATATATATATATATATATATATTTTAAATAAT | 1203 | |
| Qy | 1402 | TACCTTTATGATATATAAAATTTTGTGTTATTTTGGATTTTCATPAAGATTTATAAATCTATGTTTA | 1461 | |
| Db | 1202 | AGCTATAGAATTTTATCTTATATTTAAATGTCATTTACATAACATACAAAAATAATTAAT | 1143 | |
| Qy | 1462 | TAAATGATAATATAATTTTAAAAATAAATCTATATTAATTTCTGATTTAGTTCGATTTACCGCCT | 1521 | |
| Db | 1142 | TACTATATATATATTTTTTAAATTTATTTATATATATATATTTTGAAGCGTTATACAACTTA | 1083 | |
| Qy | 1522 | TTTTATAATTTTCAATACTGAGTAATATGAATAAATCAGTTTATCTGAAAAAGCA - AAATAAT | 1580 | |
| Db | 1082 | TATATACTATGATAAATACAGCGGATAATAACATTTTTTTTAAATATATAATTAATAATTTAT | 1023 | |
| Qy | 1581 | ATCTTTGTAAAAACAGCGTTCCGTTCAAAATGGAAGTTCAATGTGTTATTTCAATAGTTTAAATA | 1640 | |
| Db | 1022 | ATATATAATAAGTAAATATATATTTAATGTGTTCATTTATGACAACTTTAAATGTGTGTGA | 963 | |
| Qy | 1641 | TAAAGTAAAAATTTT - AAAATTAATTTGTTATTTTTTTTGTGTTTCAAAATTTAAAAATAAAATTTAT | 1698 | |
| Db | 962 | AAAATGATGATGAAAAATTTGTTATATATATATAATAAAAATAAAATTTTTTTTAAATAAT | 903 | |
| Qy | 1699 | GAGCATGGGAA 1709 | | |
| Db | 902 | TATAAATGTGCA 892 | | |

RESULT 12
US-10-204-708-84
; Sequence 84, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; TITLE OF INVENTION: by Assessing DNA Methylation
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; CURRENT FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 84
 ; LENGTH: 7664
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; US-10-204-708-84

Query Match 2.9%; Score 59.6; DB 4; Length 7664;
 Best Local Similarity 50.2%; Pred. No. 0.0016;
 Matches 204; Conservative 0; Mismatches 194; Indels 8; Gaps 2;

| | | | |
|----|------|--|------|
| Qy | 1304 | AAAAATTACTATTTTGAGAAAGTTTTGAAATGTATAGAAATTGAGTGGGACATCCATAAA | 1363 |
| Db | 408 | AATAGATTTTGATTTGTTTGGAGTATTTTGGAGAAATTTTGATTAAATAAGGTAGTTTTGAAAA | 467 |
| Qy | 1364 | AGGAAAGTGATAGAAATTTAAATCGGACAGAGGAGTAATACCTTTATGATATATATAATTT | 1423 |
| Db | 468 | AGGATTTGTAATATATTTGAAATGGAGGTGATGTGGTTTTTTTATATAAGTTTTTTTAAATTT | 527 |
| Qy | 1424 | TTGCTATTTTGATTTTC-----ATAAGATTATAAATCTATGTTTATAATGATAATAATAATTT | 1478 |
| Db | 528 | TTTTTTATATCGTTTTAGTATATTAAGAGTTTTTATTTTTTATTAATAATGTTGTAGAGTTT | 587 |
| Qy | 1479 | TAAAAATAACTATATATTAATCTCGATTAGTCGATTACCGCCTTTTATAATTTTACAATA | 1538 |
| Db | 588 | TAAGTTTTTTTTTTTCGGTTTTTTTCGTATATTATTAGTTTGAAATGTTAATTTATGTTT | 647 |
| Qy | 1539 | CTGAGTAATATGAATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAACACGCGT | 1598 |
| Db | 648 | TTTTTTTAGGTTTTGATTTAGGGGTTTGGTGATGTTCTAGTTATTTGGGTTTTTAAAGAGTAT | 707 |
| Qy | 1599 | TCGGTCAATGGGAAGTCATGTGTAATCAATAGTTTTTAATATAAAGTAAATTTTAAAT | 1658 |
| Db | 708 | TGTAATTTTA---GAGTTTAGTGTTTTTTAGATAGTAGTGTTTTTAAATTAATTTTACGA | 764 |
| Qy | 1659 | TAATTGTTATTTTTTGGTTTTCAGAAATTTAAAAATAAATTAATTGAGCAT | 1704 |
| Db | 765 | TTTTTGTATTTTAAATTAATAATTCGTATTTTTTATTTTAGAAT | 810 |

RESULT 13
 US-09-801-861-3/c
 ; Sequence 3, Application US/09801861
 ; Patent No. 6492154
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001098
 ; CURRENT APPLICATION NUMBER: US/09/801,861
 ; CURRENT FILING DATE: 2001-03-09
 ; NUMBER OF SEQ ID NOS: 10

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 53332
; TYPE: DNA
; ORGANISM: Human
US-09-801-861-3

Query Match          2.9%; Score 59.6; DB 4; Length 53332;
Best Local Similarity 47.8%; Pred. No. 0.0024;
Matches 173; Conservative 0; Mismatches 189; Indels 0; Gaps 0;

Qy 1328 TTGGAATGTATAGAAATTGAGTGGGACATCCATAAAAGGAAAGTGATAGAAATTAATGG 1387
Db 31673 TCTCAAAAAAATAAATACATATATATATATATAAAAAAATATATATAAATATAT 31614

Qy 1388 GACAGAGGGAGTAACCTTTATGATATATAAATTTTGTATTATTTGATTTCAATAAGATT 1447
Db 31613 ATAAATATATAAATAATATATAAATATAAATAAATAAATAAATAAATAAATAAATAT 31554

Qy 1448 ATAAATCTGTATATGATAAATAAATTTTAAAAATAAATACTATATAATTTCTGATTA 1507
Db 31553 ATAAATATATAAATAATATAAATAATATAAATAAATAAATAAATAAATAAATAAATAT 31494

Qy 1508 GTCGATTACCGCTTTTATATTTTACAATCTAGTAATGAGTAATGAAATCAAGTTATCTG 1567
Db 31493 TATATATAAATAATATATAAATAATATAAATAAATAAATAAATAAATAAATAAATAA 31434

Qy 1568 AAAAGCAAAATAATATCTTTGTAAACAGCGTTCGGTCAAAATGGGAAGTTCAATGTATTC 1627
Db 31433 AAAATATATAAATAATATAAATAATATAAATAAATAAATAAATAAATAAATAAATAT 31374

Qy 1628 AATAGTTTAAATAAAGTAATTTTAAATTAATTTGTTTGTTCAGAAATTTAA 1687
Db 31373 AAATATAAATACTATAAATAATATAAATAAATAAATAAATAAATAAATAAATAAATA 31314

Qy 1688 AA 1689
Db 31313 TA 31312

RESULT 15
US-10-204-708-73/c
; Sequence 73, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 73
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-73

Query Match          2.9%; Score 58.6; DB 4; Length 5152;
Best Local Similarity 49.9%; Pred. No. 0.0024;
Matches 208; Conservative 0; Mismatches 199; Indels 10; Gaps 2;

Qy 1400 AATACCTTTATGATATATAAATTTTGTATTATTTTGATTTTCAAGATTATAAATCTATGT 1459
Db 2063 AAACCTTCACCTCTATAAATATTTTATATCTCTCAAAATATAAATTTTATAAACCTTCAA 2004

Qy 1460 TATAATGATAAATAAATTTTAAAAATAAATACTATATAATTTCTGATTAGTACCGC 1519
Db 2003 AACATTAATAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1949

Qy 1520 CTTTATAAATTTTACAATCTAGTAATATGAATAAATAAATAAATAAATAAATAAATAAATA 1579
Db 1948 CCTTAATAATTTTATAAAACCAAAATCAACCAAAATAAATAAATAAATAAATAAATAAATA 1889

Qy 1580 TATCTTTGAAAACAGCGTTCGGTCAAAATGGGAAGTTCAATGTATTCATTAATTTTAAAT 1639
Db 1888 AATTTT-----ATAATCTTAAATCAAAAAAATAAATAAATAAATAAATAAATAAATAAAT 1834

Qy 1640 ATAAAGTAAATTTTAAATTAATTTGTTTATTTTCTGTTTCAAGAAATTTAAATAAATAAT 1699
```

| | | | |
|----|------|---|------|
| Db | 1833 | AAAAAAATATTTTAAATCTATAAATACTACATTTTCATAAAATTAACATAAAAAATTA | 1774 |
| Qy | 1700 | AGCATGGGAAGTTCACGGGCATCATTTGACGACACTAGACTGTTTGAACAATGTATGTCC | 1759 |
| Db | 1773 | ATTATAAAAAATTAACATAAAAAATTAATTATAATATCTTTATCCGTAAACATCTAT | 1714 |
| Qy | 1760 | GGTGTACATCTATGACCTTTTCAACTCAAACCTAGTGAATTAATGCATTTCTAGAATACAT | 1816 |
| Db | 1713 | TTTAAATCTTTATAATTTATCAATAAATACTATAAACAAAAAAATTCCTAACCAAT | 1657 |

Search completed: December 6, 2004, 22:52:26
Job time : 208.69 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 19:02:39 ; Search time 1212.45 Seconds
(without alignments)
9300.077 Million cell updates/sec

Title: US-09-806-197-1

Perfect score: 2052

Sequence: 1 catgtgtccctacagcaca.....gcaataacattcctaataatc 2052

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3695051 seqs, 2747533894 residues

Total number of hits satisfying chosen parameters: 7390102

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----------------------|-------------------|
| 1 | 99.4 | 4.8 | 8056 | US-10-473-126-386 | Sequence 386, App |
| 2 | 93 | 4.5 | 3673778 | US-10-312-841-1 | Sequence 1, Appli |
| 3 | 91.8 | 4.5 | 8056 | US-10-473-126-240 | Sequence 240, App |
| 4 | 91.6 | 4.5 | 8056 | US-10-473-126-386 | Sequence 386, App |
| 5 | 88.6 | 4.3 | 8056 | US-10-473-126-240 | Sequence 240, App |
| 6 | 83 | 4.0 | 158001 | US-10-211-179-11 | GENERAL INFORMATI |
| 7 | 75 | 3.7 | 11745 | US-10-240-453-206 | Sequence 206, App |
| 8 | 71 | 3.5 | 5930 | US-10-311-455-490 | Sequence 490, App |
| 9 | 70.6 | 3.4 | 3673778 | US-10-312-841-1 | Sequence 1, Appli |
| 10 | 70 | 3.4 | 3252 | US-10-027-632-113786 | Sequence 113786, |
| 11 | 70 | 3.4 | 3252 | US-10-027-632-113787 | Sequence 113787, |
| 12 | 70 | 3.4 | 3252 | US-10-027-632-113788 | Sequence 113788, |

| | | | | | | |
|------|------|-----|---------|----|----------------------|-------------------|
| c 13 | 70 | 3.4 | 3252 | 15 | US-10-027-632-113786 | Sequence 113786, |
| c 14 | 70 | 3.4 | 3252 | 15 | US-10-027-632-113787 | Sequence 113787, |
| c 15 | 70 | 3.4 | 3252 | 15 | US-10-027-632-113788 | Sequence 113788, |
| c 16 | 69.8 | 3.4 | 6352 | 16 | US-10-221-613-195 | Sequence 195, App |
| c 17 | 69.4 | 3.4 | 16258 | 16 | US-10-257-166-120 | Sequence 120, App |
| c 18 | 69 | 3.4 | 3673778 | 15 | US-10-312-841-2 | Sequence 2, Appli |
| c 19 | 68.6 | 3.3 | 49979 | 17 | US-10-741-601-5746 | Sequence 5746, Ap |
| c 20 | 68.2 | 3.3 | 15548 | 15 | US-10-311-455-2128 | Sequence 2128, Ap |
| c 21 | 68 | 3.3 | 11422 | 15 | US-10-311-455-191 | Sequence 191, App |
| c 22 | 68 | 3.3 | 11422 | 16 | US-10-257-166-17 | Sequence 17, Appl |
| c 23 | 67.8 | 3.3 | 15732 | 14 | US-10-239-676-95 | Sequence 95, Appl |
| c 24 | 67.8 | 3.3 | 15732 | 15 | US-10-240-453-107 | Sequence 107, App |
| c 25 | 67.6 | 3.3 | 5984 | 17 | US-10-433-793-24 | Sequence 24, Appl |
| c 26 | 67.6 | 3.3 | 6050 | 15 | US-10-311-455-1984 | Sequence 1984, Ap |
| c 27 | 67.4 | 3.3 | 5748 | 15 | US-10-311-455-1115 | Sequence 1115, Ap |
| c 28 | 67.4 | 3.3 | 18154 | 15 | US-10-311-455-228 | Sequence 228, App |
| c 29 | 67.2 | 3.3 | 5807 | 15 | US-10-311-455-1128 | Sequence 1128, Ap |
| c 30 | 67 | 3.3 | 2053 | 13 | US-10-027-632-97533 | Sequence 97533, A |
| c 31 | 67 | 3.3 | 2053 | 15 | US-10-027-632-97533 | Sequence 97533, A |
| c 32 | 66.6 | 3.2 | 419 | 9 | US-09-960-352-11234 | Sequence 11234, A |
| c 33 | 66.6 | 3.2 | 6079 | 15 | US-10-311-455-394 | Sequence 394, App |
| c 34 | 66.6 | 3.2 | 6292 | 16 | US-10-221-714A-461 | Sequence 461, App |
| c 35 | 66.4 | 3.2 | 18624 | 15 | US-10-311-455-1676 | Sequence 1676, Ap |
| c 36 | 66.2 | 3.2 | 1501 | 18 | US-10-473-126-328 | Sequence 328, App |
| c 37 | 66.2 | 3.2 | 3252 | 13 | US-10-027-632-113786 | Sequence 113786, |
| c 38 | 66.2 | 3.2 | 3252 | 13 | US-10-027-632-113787 | Sequence 113787, |
| c 39 | 66.2 | 3.2 | 3252 | 13 | US-10-027-632-113788 | Sequence 113788, |
| c 40 | 66.2 | 3.2 | 3252 | 15 | US-10-027-632-113786 | Sequence 113786, |
| c 41 | 66.2 | 3.2 | 3252 | 15 | US-10-027-632-113787 | Sequence 113787, |
| c 42 | 66.2 | 3.2 | 3252 | 15 | US-10-027-632-113788 | Sequence 113788, |
| c 43 | 66.2 | 3.2 | 6175 | 15 | US-10-311-455-1280 | Sequence 1280, Ap |
| c 44 | 65.8 | 3.2 | 15161 | 16 | US-10-221-613-386 | Sequence 386, App |
| c 45 | 65.2 | 3.2 | 34722 | 17 | US-10-322-281-700 | Sequence 700, App |

ALIGNMENTS

RESULT 1

US-10-473-126-386

; Sequence 386, Application US/10473126

; Publication No. US20040234973A1

; GENERAL INFORMATION:

; APPLICANT: Epigenomics AG

; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell

; TITLE OF INVENTION: proliferative disorders

; FILE REFERENCE:

; CURRENT APPLICATION NUMBER: US/10/473,126

; CURRENT FILING DATE: 2003-09-26

; NUMBER OF SEQ ID NOS: 1258

; SEQ ID NO 386

; LENGTH: 8056

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-473-126-386

Query Match 4.8%; Score 99.4; DB 18; Length 8056;

Best Local Similarity 44.5%; Pred. No. 9.7e-08;

Matches 624; Conservative 0; Mismatches 761; Indels 18; Gaps 5;

| | | | |
|----|------|---|------|
| QY | 548 | ATGCTTCTCAAAATGTTTTTATATGTAATAAATGCCCCATCCAGGATAAGTAATAATTC | 607 |
| DB | 1975 | ATTAAATTAATATTTTATTTAAATTTTAAATAATTTTAAATATATTTAAATAATAATAT | 2034 |
| QY | 608 | CGGTTTAACCGATGTTGTTAATATATATGTTACACTTACAAGGATATTCGTAATACCT | 667 |
| DB | 2035 | ATTTTAAATATTTTATTTTAAATAATTTTAAATAATTTTAAATAATATTAATATTAAT | 2094 |
| QY | 668 | TTAGACGACAGAGACTTAGTGCAAAAATGGCGGTGTAACAGCCTAGACTTGCTCAC | 727 |
| DB | 2095 | AATAATTATAAATTTTAAATAATTTTAAATAATTTTAAATAATTTTAAATAATAAAAA | 2154 |

| Query Match | 4.3% | Score 88.6 | DB 18 | Length 8056 |
|-----------------------|----------------|--|-----------|-------------|
| Best Local Similarity | 45.4% | Pred. No. 8.2e-06 | | |
| Matches | Conservative 0 | Mismatches 519 | Indels 12 | Gaps 3 |
| Qy | 736 | AGATAATGGTTAGTATATAATAGTAGGATCTACAACTGACATTAATAATTTAGAGCTATTAAAT | 795 | |
| Db | 1399 | ATATATTTTTTATATAAAAAAATAAATTTTTTATATAAAAAATTTAAATTTTAAATTTATTTT | 1340 | |
| Qy | 796 | TAAAGTTACTAATAATAAAGAGAGGTTAGTAAACAGAAAGCAGGCTAAACAAAGAGCTTGC | 855 | |
| Db | 1339 | TTATTAAAAAATAATAATAATTTTTTTTATAAATTTATTAAATTTTAAAAAATAATTCAGTTT | 1280 | |
| Qy | 856 | TGCTGTGCTGTTAGTCTGTGAGCTCATTTCTTTAAAGTAATGTAAACCTGATCTAAAG | 915 | |
| Db | 1279 | TTATATATTAATAATAATAGCTGACGTTCTTTTTTTTTTTTCGAAAAAATTTATTTATTATTA | 1220 | |
| Qy | 916 | CACATAGAAAATTTAGTACAGGTTAAAACTTTTTACAAGAAATTTATTTAAAAACGAAAAATCAT | 975 | |
| Db | 1219 | AAAAAATTAATAATAATAATTAATAAATTTTTTAAAAATAAATTAATAATTCGAAAAATATA | 1160 | |
| Qy | 976 | TTTATACATGCTCTCGGCTGTCATTTATTAATAGGGATCACTTACTGATCATCCATTAAA | 1035 | |
| Db | 1159 | TTAATTTTAAATTTTCGAATATCGAAAAAATAATTAAAAAACGAATATATAACGTTTTTAA | 1100 | |
| Qy | 1036 | ACCTTGTGTAAAAACAAATTCATAGAGATAAAATATCTTACAATGAAAAAGAGCAATGTC | 1095 | |
| Db | 1099 | AACGTTAAAAAATAAACAAGATTAATAACGAAATTTTTTTTATTTTACGTTAAATTAATTCGA | 1040 | |
| Qy | 1096 | TCCTTTGAAAAAACAATPAGGTACTCTCCCTCGCTCCCTCGAAATGTATACATATGGATTTG | 1155 | |
| Db | 1039 | AAATTAATAAATAATAAAAAAATTTTTTAAAAATAATTTATTTTAAAAAATAATATAATATA | 980 | |
| Qy | 1156 | ACCGGAGACTAGAAAAATCTATAAAGTAATGTAGAGTAAAGAGAGAGAGAGAGAAAG | 1215 | |
| Db | 979 | TTTTTATTCGTTTAAATAAAAAATAAATAATTTATTTTTTAAAAAATAATTTTATTCGAAAAAC | 920 | |
| Qy | 1216 | TGGTAAAGTAGCGGGACCCACCAATATATAAATTTGATAGATTTTTAGAAAAAGTAGTGAAG | 1275 | |
| Db | 919 | GACGAAAAATTAATAAATAATAAATTTTAAATAATTTACGAAATTTTATACGTAAAAAATAA | 860 | |
| Qy | 1276 | TAGTGGGTGGTGGGATTTTTTATATATAAAAAATTTTACTAT - TTTGAGAAAGTTTTCGAAA | 1334 | |
| Db | 859 | TAAAAAATAACGATTTTTTTTTTAAATAAATAATTTTAAATAATCGTTTATAAATTTATTTTA | 800 | |
| Qy | 1335 | TGTATAGAAATTTAGTGGGCATCCATAAAGGAAAGTGTATAGAAATTTAAATGGACAGAG | 1394 | |
| Db | 799 | AAATATATAATTAATAATAAATTTTATAATTTTATAATTTTAAATTTTAAATTTTAAATTTAAT | 740 | |
| Qy | 1395 | GGAGTAAATCCTTTATGATATAATAAATTTTTTGGTTAT - - - - - TTTGATTTCAATAAGTTA | 1448 | |
| Db | 739 | TATAATAATAAATTTATTTTAAATTTAAATTTTTTAAATATAAAAAAATTTTTTTTTATATTTA | 680 | |
| Qy | 1449 | TAAATCTATGTTATAAGATAATAATAATTTTTTAAAAATAAATACTATATAATTTCTGATTAG | 1508 | |
| Db | 679 | TTTTTTTTTTTTTTTTTAAATTAATTTATTTTAAATAATTTTTTTTAAAAAATTTTTTAAATTT | 620 | |
| Qy | 1509 | TCGATTACCGCCTTTTATAA - - - - - TTTTACAATCTAGTAAATATGAATAAATCAGTTA | 1563 | |
| Db | 619 | TATATAATAATTTTTTTTATAAATAATAATTTTAAATAATAAATAATTTTAAATAAACGTTTA | 560 | |
| Qy | 1564 | TCGAAAAGCAATTAATATCTTTGTAAAAACAGCGTTCGGTCAATGGGAAGTTCAATGTT | 1623 | |
| Db | 559 | TAATAATTTATATAAATAATTTTAAATTTTAAATATAATAATTTTAAATATAATAATAATTA | 500 | |
| Qy | 1624 | ATTCAAATAGTTTTTAATAATAAAGTAAATTTTTTAAATTAATTTGTTATTTTCTGTTTTCAGAAAT | 1683 | |
| Db | 499 | TTTTTAAATTTTTTTTATTAATAATAATTTTATTTTAAATTAATTTTAAAAATTAATTTTATTTT | 440 | |
| Qy | 1684 | TTAAAAATAAATA 1696 | | |
| Db | 439 | TTAAAAAATAA 427 | | |

RESULT 6
 US-10-211-179-11
 ; GENERAL INFORMATION:
 ; APPLICANT: Nicholas M. Dean
 ; APPLICANT: Kenneth W. Dobie
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOTYROSYL PHOSPHATASE ACTIVATOR EXP
 ; FILE REFERENCE: PTS-0011
 ; CURRENT APPLICATION NUMBER: US/10/211.179
 ; CURRENT FILING DATE: 2002-08-01
 ; NUMBER OF SEQ ID NOS: 119
 ; OTHER INFORMATION: n = a, t, c, or g
 US-10-211-179-11

Query Match 4.0%; Score 83; DB 16; Length 158001;
 Best Local Similarity 46.0%; Pred. No. 0.00023;
 Matches 458; Conservative 0; Mismatches 530; Indels 7; Gaps 5;

| | | | |
|----|--------|---|--------|
| Qy | 708 | AACAGCCTGACCTGGTGCATCTGATAAATAGATAATCTTACTAGTATAATATAGTAGAATCTA | 767 |
| Db | 126953 | AAATATATTATATATATATATAATAAATATATATATATATATATATAATAATAATATATATA | 127012 |
| Qy | 768 | CAATGACATTAATAATTAGAGCTATTAATTAAGTTACTAAATAAATAAGAGAGGTTAGTAAA | 827 |
| Db | 127013 | ATAATATATAATATAATAATAATATATATAATTTGTAATATAATAATAATAATAATAA | 127072 |
| Qy | 828 | CAGAAAGCAGGTAAAAACAAGAGCTTCTGCTGTGCTTTAGTTGTTGTGAGCTCATTTTC | 887 |
| Db | 127073 | AAAAATAATATAAATATATAAAAAATATAAATATATATATATATAATAAATATATAAAAAAT | 127132 |
| Qy | 888 | TTTAAAAAGTAAATGTAAGTAACTGCTTAAAGCACATAGA-AAATTAGTACAGGTTTAAACCTTT | 946 |
| Db | 127133 | ATAATATATATTTATATAAATAATAAAAAATATAAATAATATAAATAATATATATATAAATAATA | 127192 |
| Qy | 947 | TACAAGAAATTTATATTAACGAAATCAATTTTATAACATGCTCTCGGCTGTCTCAATTATA | 1006 |
| Db | 127193 | TAAAAATATAAATATATATTTATATAAATAATAAAAAATATAAATAATATAATATATATATA | 127252 |
| Qy | 1007 | TAGGGATCACTTACTGATCATCCATTTAAACCTTGTTTAAACAAATTTCAATGAGATAAA | 1066 |
| Db | 127253 | AAATATAAAAAATATAAATATATATATTTATATAAATATAAAAAATATAAATAATATAAT | 127312 |
| Qy | 1067 | TATCTTACAATG-AAAAAGAAGGCAATGTCTCTTTTGAAAAAACAATAAGTACTCCCTCC | 1125 |
| Db | 127313 | TATATAAATATATAAATATATAATAATATAATAATATAAATAATAAATAATAATAATAT | 127372 |
| Qy | 1126 | GTCCCTCTGAAATGTATACATATGGATTGGACACGGAGACTTAAGAAAAATGTATAAAGTA | 1185 |
| Db | 127373 | ATATTATATAAATATAAATAATAAATAATAATAATATATATATATAAATAATAAATAATA | 127432 |
| Qy | 1186 | ATGTAGAGTAAAGAAGAAAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATAT | 1245 |
| Db | 127433 | TAATATATATTTATATATAAATAATAAATAATAAATAATAAATAATAAATAATAAATAAT | 127492 |
| Qy | 1246 | AAITGATAGATTTTGAAGAAAGTAGTTGAAAGTAGTGGGTGGGATTTTATATATTATAA | 1305 |
| Db | 127493 | AAATATAAATAATAATAATAATAATAATAAATAATAATAAATAATAATAATAATAATAAT | 127552 |
| Qy | 1306 | AAATTTTACTATTTTGAGAAAGTTTGGAAATGTATAGAAATTTGAGTGGGACATCCCAATAAG | 1365 |
| Db | 127553 | AAAAAAATATACAATATAATAATATAAATAATAAATAATAAATAATAAATAATAAATAATA | 127612 |
| Qy | 1366 | GAAAGTGTATAGATTAATTTGGGACAGAGGAGTA-ATACCTTTTATGATATATAAATTTT | 1424 |
| Db | 127613 | TATATAATATATAAAAAAATAAATAATAATAATAATAATAATAATAATAATAATAATAAT | 127672 |
| Qy | 1425 | TGTTATTTTTCATTTTATAAGATTATAAATCT-ATGTTTATAATGATAATAAATTTTAAAA | 1483 |
| Db | 127673 | AAAAATATAAATATAATAATAATAATAATAAATAATAAATAATAAATAATAATAATAATA | 127732 |
| Qy | 1484 | ATAATACTATATAATTTCTGATTTAGTCGATTTACCGCCTTTTATAATTTTACAATCTGAG | 1543 |
| Db | 127733 | TATATAATATAATAATAAAAAATAAATAATAATAATAATAATAATAATAATAATAATAAT | 127792 |


```
Db 1714513 TATATATTTTATATATATTTTATATATATATATTTTATATATATATTTTATATATATATTTT 1714572
QY 806 AFAAATAAGAGAGGTTAGTAAGACAGAAAGCAGGTAAACAGAGAGCTTGCCTGTGTGT 865
Db 1714573 ATATATATATATTTTATATATATATTTTATATATATATATTTTATATATATATTTTATAT 1714632
QY 866 TTAGTTGTGT-GAGCTCATTTCTTTAAAGTAATGTAAACTGATCTAAAGCACATAGAA 924
Db 1714633 ATATATATTTTATATATATTTTATATATATATTTTATATATATTTTATATATATTTT 1714692
QY 925 ATTTAGTACAGGTAAACCTTTTCAAGAAATTTTATTAATAACGAAACATTTTATACA 984
Db 1714693 TATATGTATATTTTATATGTATATTTTATATGTATATTTTATATGTATATTTTATATATA 1714752
QY 985 TCTCTCGGCTGTCATTATAATAGGATCACTTACTGATCATCCATTAAACCTTTGTTA 1044
Db 1714753 TATTTTATATGTATATTTTATATATATATTTTATATATATTTTATATATATATTTTATAT 1714812
QY 1045 AAACAATTCATGAGATAAATATCTTACAATGAAAGAGGACAATGTCTCTTTGAAA 1104
Db 1714813 ATATATTTTATATATTTTATATATATATAGTATATATATTTTATATGTATATTTTATATA 1714872
QY 1105 AAACAAATAGTACTCCCTCCCTCTGAAATGTATACATATGATTTGGACACGGAGA 1164
Db 1714873 TATTTTATATATATATTTTATATATATATTTTATATATATTTTATATATATATTTTGT 1714932
QY 1165 CTGAGAAAATGTAATAAGTAATAGTAGAGTAAGAAAGAGAGAAAGGTTGGTAAAG 1224
Db 1714933 ATATATATATTTTGTATATATATATTTTATATATATATTTTATATATATATATTTTAT 1714992
QY 1225 TAGCGGGACCCACCAATATATATTTGATAGATTTAGAAAAGTAGTTGAAAGTAGGGTG 1284
Db 1714993 ATATATATATATTTTATATATATATTTTATATATATATA----TATTTTATATATATATTT 1715048
QY 1285 GGTGGGATTTTATATATAAATAATTTACTATTTTGGAGAAAGTTTGGAAATGTATAGAAT 1344
Db 1715049 ATATATATATATATTTTATATATATATTTTATATATATATTTTATATATATATATTT 1715108
QY 1345 TGAGTGGGACATCCATAAAGGAAAGTGATAGAATTAAGTGGGACAGAGGGAGTAATAC 1404
Db 1715109 TTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATAT 1715168
QY 1405 CTTTATGATATAAATTTTGTATTTTGTATTTTATATATATATATTTTATATATATATATTT 1464
Db 1715169 ATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTT 1715228
QY 1465 TGATAATATAATTTTAAATAATACTATATATTAATTTCTGATTAGTCGATTACCGGCTTTT 1524
Db 1715229 TTTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATATA 1715288
QY 1525 AFAATTTTACAATCTGAGTAATAGAAATAATCAAGTATCTATCGAAAAGCAATAATATCT 1584
Db 1715289 TATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATATATATATTTTATA 1715347
QY 1585 TTGTAACACAGGTTCCGTCAAATGGGAGTTCAGTGTATTCATAGTTTAAATATAAA 1644
Db 1715348 TTTATATGTATATTTTATATATATATATTTTATATGTATATTTTATATATATATTTTATATA 1715407
QY 1645 AGTAAATTTTAAATTAATTTGTTATTTT 1671
Db 1715408 TGATATTTTATATGTATATTTTATAT 1715434
```

RESULT 10
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20020198371a1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827.129

```
; CURRENT APPLICATION NUMBER: US/10/027,632  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 113786  
; LENGTH: 3252  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-113786  
  
Query Match 3.4%; Score 70; DB 13; Length 3252;  
Best Local Similarity 43.6%; Pred. No. 0.012;  
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;  
  
QY 621 TTGTTAATATATATATGTTTACACTTACAGAGGATATTCGTAACTCTTTAGACGACAAGA 680  
Db 2314 TTAATCTACAAATTTTATACTTATATATGATTTAATATGTTTATATCTGTAATATA 2255  
  
QY 681 GACTTAGGTCAAAAATGGACGCTCGTAAACAGCCTAGACTTGGTCACGTAGATAAATAGATA 740  
Db 2254 TATTTAACTGTGTATGGTAACATATGAATAGTTAATATATTTAAATTTAATATAAATAC 2195  
  
QY 741 ATTGTTAGTATAATATAGTAGGATCTACAATGACATTAATAATTAGAGCTATTAATTAAGT 800  
Db 2194 ATTATAATATAAATTTTAAACAATATATCATTAATACATATTCATATTTATATATTTTAA 2135  
  
QY 801 TACTTA-----ATAAATAAGAGAGGTTAGTAAACAGAAAGCAGGTAATAAACCAAG 848  
Db 2134 TATAATACCATAATACTTATATTTAAAGCTAATTTAATATATATACATTAATACTAATA 2075  
  
QY 849 AGCTGTGCTGTGTGTTTGTGAGCTCATTTCTTTAAAGTAATGTAAACTGA 908  
Db 2074 ATCTTTATACCTTATTTGTTAACTTACATTTAAATTAATTAATATTTAAATTAACCTTAT 2015  
  
QY 909 TCTAAAGCACATAGAAATTTAGTACAGGTTA-AAACCTTTTACAGAAATTTATATTTAAACG 967  
Db 2014 TCTGGTTAACTATTATATATAGTTAAACCATATTTAAATATATATATATATAAATAATACA 1955  
  
QY 968 AAAATCACTTTTAAACATGTCCTCGGCTGTCAATTAATAAGGATCACTTACTGATCAT 1027  
Db 1954 TGAATATATATATATAAATTTTGCATTTTATTTATATATATATATATATATTTAATAT 1895  
  
QY 1028 CCATTAACAACTGTTTAAACAAATTCATGAGATATAAATATCTTACAAATGAAAGAGAGG 1087  
Db 1894 ATAATATTTATATATATATATTTTAAATAAACAATATTTTATATATATATATATTTCAATA 1835  
  
QY 1088 ACAATGCTCTCTTGAATAAACAATAATAGTACTCCCTCCGCTCGGCTGAAATGTATACATA 1147  
Db 1834 AATATTTATATATATATATTTTAAATAAACAATAATATATATATATATTTTAAATATACAAT 1775  
  
QY 1148 TGGATTGGACACGGAGCTAAAGAAAATGTATAAGTAATGTAGAGTAAAAAGAAAGAGA 1207  
Db 1774 TATATATTTATATATATTTTAAATATACAATTTTATATATATATATATATTTTATATACA 1715  
  
QY 1208 AAGAAAGTGGGTAAAGTAGCGGGACCCACCAATATATAATTTGATAGATTAGAGAAAGTA 1267  
Db 1714 TTTATATATATATATATTTTAAATATACAATAATTTTATATATATATATATTTTAAATATA 1655  
  
QY 1268 GTTGAAAGTAGTGGGTGGGATTTTATATATATAAATAATTTTACTATTTTGTGAGAAAGT 1327
```



```
RESULT 12
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 32520
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113788
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113786

Query Match      3.4%; Score 70; DB 13; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

QY 621 TTGTTAATATATATGTTTACACTTACAGAGATATTCGTATATCTTTTAGACGACAGA 680
DB 2314 TTAATACTAACAAATTTTATACCTTATATTAATGAGTTTAATATGTTTACTGTAATATA 2255

QY 681 GACTTAGTCAAAAATGAGCGCTGGTAAACAGCCTAGACTTGGTCACATGATAAATAGATA 740
DB 2254 TATTTAACTGTGTATGGTAACATATGTAATGTTAATATATTAATTAATATATATAC 2195

QY 741 ATTGTTAGTATATATAGTAGGATCTACAAATGACATTAATAATAGAGCTATTAATTAAGT 800
DB 2194 ATTATAATATAAAATTTATAACAATATATCATATACATAAATTCATATATATATTTAATA 2135

QY 801 TACTA-----ATAATAAGAGGTTAGTAACAGAAAGCAGGTAACAAACAG 848
DB 2134 TATAATACCATAAATACTTATATTAATAAGCTAAATTTAATATATTTACATTAATCTAATA 2075

QY 849 AGCTTGCTGCTGTGTTAGTTGTGTGAGCTCAATTTCTTTAAAGTAATGTAACCTGA 908
DB 2074 ATCTTTATCTTATGTTTAACTTACATTAATTAATTAATTAATTAATTAATTAATTAAT 2015

QY 909 TCTAAGCACAAGAAATTTAGTACAGGTTA-AACTTTTACAGAAATTTATPATTAACG 967
DB 2014 TCTGGTTAACTATATATATAGTTAAACCATATTAATAATATAAATAATATAAATAATA 1955

QY 968 AAAATCATTTTATAACATGCTCTCGGCTGTCATTAATTAATAGGATCACTTACTATCAT 1027
DB 1954 TGATATATTTATATATATAATTTTGCATTTTATTTATTTATATATATATATATTTAATAT 1895

QY 1028 CCATTAAACACCTTGTAAACAAAATTCATAGAGATAAAATCTTACAAATGAAAAGAGG 1087
DB 1894 ATAATATTTATATATATATATTTTAATAACAATATTTATATATATATATATTTCAATAAC 1835

QY 1088 ACAATGCTCTTTTGAATAAACAATAAGTAGTACTCCCTCCGCTCGCTGAAATGTATACATA 1147
DB 1834 AATATTTATATATATATATTTTAATAAACAATATATATATATATATATTTAATATACATAT 1775
```

```
QY 1148 TGGATTGGACACGGAGACTAAGAAAATGTATATAAGTAATGTAGNGTAAAAAGAAAGAGA 1207
DB 1774 TATATATTATATATATATTTTAAATATACAATATTTATATATATATATATATATACAATA 1715
QY 1208 AAGAAAAGTGGGTAAAGTAGCGGACCACCAATATATAAATTTGATAGATTTTAGAAAAGTA 1267
DB 1714 TTTATATATATATATATTTTAAATATACAAATATTTATATATATATATATATATATACAA 1655
QY 1268 GTTGAAGTAGTGGTGGTGGGATTTTATATATATAAAAAATTTACTATTTTGGAGAAAGT 1327
DB 1654 TATTTATATATATATATATTTTAAATATACAATATTTATATATATATATATATATATATA 1595
QY 1328 TTTGAATGTATAGAAATTTGAGTGGGACATCCATAAAGGAAAGTGTATAGATTTAAATGG 1387
DB 1594 TTTATATATATATATAATTTGATTTGATCATACATAAATTTATATATCAATATATTTCAATA 1535
QY 1388 GACAGAGGAGTAATACCTTTTATGATATATAAAATTTTGTATTTTGTATTTTCATAAGATT 1447
DB 1534 TTAATATATGAACTACTATTAATATATATATATATTTGCTATTTCTTATATATATAAGA 1475
QY 1448 ATAAATCTATCTTATATAATGATA-ATATAATTTTAAAAATATACTATATATTAATTCGATT 1506
DB 1474 ATTATATATATATATTTGATTTATATATATATATATATATATATATATATATATATAT 1415
QY 1507 AGTCGATTACCGCTTTTATATATTTTACAAATCTGAGTAATATGATTAATCAAGTTACT 1566
DB 1414 TATTATGTAATTTGATTTTATTTATTTATATATAAATTAATGTAATGCAAAATTTATATAATA 1355
QY 1567 GAAAAGCAAAATAATATCTTTGTAATAACAGCGCTCGGTCAAAATGGGAAGTTCATGTGTAAT 1626
DB 1354 TTATATAATATATAATATATATATTTATATATATATATATATATATATATATATATATA 1295
QY 1627 CAATAGTTTTATATAAAGTAATTTTAAATTAATTTGTTATTTTGTTCAGAAATTTA 1686
DB 1294 TATAATAGTTAGCATATTAATAAGTTAATTTACTATATATTAATTAATTAATTAAGTTAA 1235
QY 1687 AAATAAAT 1694
DB 1234 CAGTAAGT 1227
```

```
RESULT 13
US-10-027-632-113786/c
; Sequence 113786, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 32520
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113786
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113786
```

Query Match 3.4%; Score 70; DB 15; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

QY 621 TTGTTAATATATATGTTTACACTTACAGAGGATATTCGTAATACCTTTTAGACGACAAGA 680
DB 2314 TTAATACTACAAATTTTATACATATATTAATGAGTTTAATATGTTTATACCTGTAATATA 2255

QY 681 GACTTAGTCAAAAATGGAGCGTGGTAAACAGCCTAGACTTGGCTCACTGATAAATAGATA 740
DB 2254 TATTTAACTGTGTATGTTACATATGATGAATAGTTAATATATTAATTAATATATATAC 2195

QY 741 ATTGTTAGTATAATATAGTAGGATCTCAATGATGATTAATAAATTTAGAGCTATTAAATTAAGT 800
DB 2194 ATTATAATATAAATTTTAAACAATATATCATATACATAATTCATATATATATATTTAATA 2135

QY 801 TACTA-----ATAAATAAGAGAGGTTAGTAAACAGAAACAGGTAACAAACAG 848
DB 2134 TATAATACCATAAATACTTATATTAATAAGCTAAATTTAATATATATTAATTAATTAATA 2075

QY 849 AGCTTCGCTGCTGTTAGTTTGGAGCTCATTTCTTTAAAGTAATGTAACTGCA 908
DB 2074 ATCTTTTACTTATTTGTTAACTTACATTAATAATTAATTAATTAATTAATTAATTAATTA 2015

QY 909 TCTAAAGCACATPAGAAATTTTAGTACAGGTTA-AAACCTTTTACAAGAAATTTATATTAACG 967
DB 2014 TCTGGTTAACTTATATATATAGTTAAACCATATTTAAATATATATAAATATATAAATAACA 1955

QY 968 AAAATCATTTTATAACATGCTCTCGGCTGCTCATTTAATAAGGATCACTTACTGATCAT 1027
DB 1354 TGTATATATATATATATATATTTTGCATTTTATTTATTTATATATATATATATATTTAATAT 1295

QY 1627 CAATAGTTTTAAATATAAAGTAAATTTTAAATTTAAATTTGTTATTTTGTGTTTTCAGAAATTTA 1686
DB 1294 TATAATAGTTAGCATATTAATAAGTTAAATTTTACTATATTAATTAATTAATTAATTAAGTTAA 1235

QY 1687 AAATAAAT 1694
DB 1234 CAGTAAGT 1227

RESULT 14
US-10-027-632-113787/c
; Sequence 113787, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108927.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 32520
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113787
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113787

Query Match 3.4%; Score 70; DB 15; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;

QY 621 TTGTTAATATATATGTTTACACTTACAGAGGATATTCGTAATACCTTTTAGACGACAAGA 680
DB 2314 TTAATACTACAAATTTTATACATATATTAATGAGTTTAATATGTTTATACCTGTAATATA 2255

QY 681 GACTTAGTCAAAAATGGAGCGTGGTAAACAGCCTAGACTTGGCTCACTGATAAATAGATA 740
DB 2254 TATTTAACTGTGTATGTTAAACATATGAAATAGTTAATATTAATTAATATATATAC 2195

QY 741 ATTGTTAGTATAATATAGTAGGATCTCAATGATGATTAATAAATTTAGAGCTATTAAATTAAGT 800
DB 2194 ATTATAATATAAATTTTAAACAATATATCATATACATAATTTCAATATATATATTTAATA 2135

QY 801 TACTA-----ATAAATAAGAGAGGTTAGTAAACAGAAACAGGTAACAAACAG 848
DB 2134 TATAATACCATAAATACTTATATTAATAAGCTAAATTTAATATATATTAATTAATTAATA 2075

QY 849 AGCTTCGCTGCTGTTAGTTTGGAGCTCATTTCTTTAAAGTAATGTAACTGCA 908
DB 2074 ATCTTTTACTTATTTGTTAACTTACATTAATAATTAATTAATTAATTAATTAATTAATTA 2015

QY 909 TCTAAAGCACATPAGAAATTTTAGTACAGGTTA-AAACCTTTTACAAGAAATTTATATTAACG 967
DB 2014 TCTGGTTAACTTATATATATAGTTAAACCATATTTAAATATATATAAATATATAAATAACA 1955

QY 968 AAAATCATTTTATAACATGCTCTCGGCTGCTCATTTAATAAGGATCACTTACTGATCAT 1027
DB 1954 TGTATATATATATATATATTTTGCATTTTATTTATTTATATATATATATATTTAATAT 1895

1028 CCATTAAACCTTGTGTTAAACAAATTCATAGAGATAAAATATCTTACCAATGAAAAGAGG 1087
Db |||||
1894 ATAATATTATATATATATATTTAATAACCAATATTTATATATATATATATATATCAATAAAC 1835
Qy |||||
1088 ACAATGCTCTTGTGAAAACAAATAGGTACTCCCTCCGTCCTCGAATGTATACATA 1147
Db |||||
1834 AATATTTATATATATATATTTATTAATAACAATATATATATATATATATATATATATATAT 1775
Qy |||||
1148 TGGATTGGACACGGAGACTAAGAAAATGTATATAAGTAGTAAAAAGAAAAGAGA 1207
Db |||||
1774 TATATATATATATATATTTAATATACATATATTTATATATATATATATATATATATATACATA 1715
Qy |||||
1208 AAGAAAGTGGTAAAGTAGCGGGACCCACCAATATATAATTTGATAGATAGATTTAGAAAAGTA 1267
Db |||||
1714 TTTATATATATATATATTTAATATACAAATTTATATATATATATATATATATATATATACAA 1655
Qy |||||
1268 GTTGAAGTAGTGGTGGGATTTTATATATATAAAATTTTACTATTTTGAGAAAGT 1327
Db |||||
1654 TATTTATATATATATATTTAATATACAAATTTTATATATATATATATATATATATATATATA 1595
Qy |||||
1328 TTTGAAATGTATAGAAATGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAATTAATATGG 1387
Db |||||
1594 TTTATATATATATATATTTGATGATCATACATAATTTATATATATATATATATATATATATAT 1535
Qy |||||
1388 GACAGAGGAGTAATACCTTTATGATATATAAAATTTTGTATTTGATTTGATTCATAGATTT 1447
Db |||||
1534 TTAATATATGAACTATTAATA 1475
Qy |||||
1448 ATAAATCTATGTTATATATGATA-ATATAATTTTAAATAATATACATATATATATATCTGATT 1506
Db |||||
1474 ATTATATATATATGTTAT 1415
Qy |||||
1507 AGTCGATTACCGCTTTTATATATTTTCAATATCTGAGTAATATGAATAAATCAAGTTATCT 1566
Db |||||
1414 TATTATGTAATGATTTTATATGTTATATATATATATATATATATATATATATATATATATA 1355
Qy |||||
1567 GAAAGCAATATATATCTTTGTAACACAGCGTTCGGTCAATGGAAGTTCATGTTAT 1626
Db |||||
1354 TTTAT 1295
Qy |||||
1627 CAATAGTTTATATATAAAGTAAATTTTAAATTAATGTTATTTTGTTCAGAAATTTA 1686
Db |||||
1294 TATAATAGTAGCAATTAATTAAGTTAATTTACTATATATTAATTAATTAATTAAGTTAA 1235
Qy |||||
1687 AAATAAAT 1694
Db |||||
1234 CAGTAAGT 1227

RESULT 15

US-10-027-632-113788/c
; Sequence 113788, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 113788
; LENGTH: 3252
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-113788

Query Match 3.4%; Score 70; DB 15; Length 3252;
Best Local Similarity 43.6%; Pred. No. 0.012;
Matches 474; Conservative 0; Mismatches 600; Indels 14; Gaps 3;
Qy 621 TTGTTAATATATATGTTTACATTTCAAGAGGATATTCGTAAATCTTTTACGCGCAAGA 680
Db |||||
2314 TTAATACTAACAAATTTTATATCTTATTAATAGTTTAAATATGTTTATCTGTAATATA 2255
Qy 681 GACTTAGGTCAAAATGGACGCTGGTAAACAGCCTAGACCTTGGTCACGTAAATAGATA 740
Db |||||
2254 TATTTAACTGTGTATGTTAACATATGAATAGTTAATATATTAATTTAATATATAATAC 2195
Qy 741 ATTGTTAGTATAATATAGTAGGATCTCAATGACATTAATAAATTAGAGCTATTAATTAAGT 800
Db |||||
2194 ATTATATATATAATTTTATAACAATATATCATATACATAATTCATATATATATATTTAATA 2135
Qy 801 TACTA-----ATAAATAAGAGAGGTTTAGTAAACAGAGCGGTAAACAAAG 848
Db |||||
2134 TATAATACCAATAAATCTTATATTTAAAAAGCTAATTTAAATATATATACATTAATACTAATA 2075
Qy 849 AGCTTGCTGCTGCTGTTAGTTGTTGAGCTCATTTCTTTAAAGTAATGTAACGTA 908
Db |||||
2074 ATCTTTTACTTATTTGTTAACTTACATTAATAATTAATTAATATATATAATTAACCTTATTA 2015
Qy 909 TCTAAGCACATACAAATTTTAGTACAGGTTA-AACTTTTACAGAAATTTATATTTAAACG 967
Db |||||
2014 TCTGTTAACTTATATATATAGTTAACATATTAATAATATATATATATATATATATATATA 1955
Qy 968 AAAATCATTTTATTAACATGCTCTCGGCTGTCTTAAATAAGGATCACTTACTGATCAT 1027
Db |||||
1954 TGATATATATATATAAATTTGCAATTTTATTTATATATATATATATATATATATATATAT 1895
Qy 1028 CCATTAACACCTTCTTAAACAAATTCATGAGATAAAATCTTACATGAAAGAGG 1087
Db |||||
1894 ATAATATATATATATATATTTTAATAACAATATTTATATATATATATATATATATATAT 1835
Qy 1088 ACAATGCTCTTTGAAAAACAAATAGTACTCCCTCCGCTCGCTGAAATGTATACATA 1147
Db |||||
1834 AATATTTATATATATATATTTTAAATAACAATATATATATATATATATATATATATATAT 1775
Qy 1148 TGGATTGGACACGGAGACTAAGAAAATGTATATAAGTAATGTAGAGTAAAAAGAAAGAGA 1207
Db |||||
1774 TATATATATATATATTTAATATACAAATTTTATATATATATATATATATATATATATATACA 1715
Qy 1208 AAGAAAGTGGTAAAGTAGCGGGACCCACCAATATATATATTTGATAGATTTAGAAAAGTA 1267
Db |||||
1714 TTTATATATATATATATTTTAAATATACAAATTTTATATATATATATATATATATATATACAA 1655
Qy 1268 GTTGAAGTAGTGGTGGGATTTTATATATATAAAATTTTACTATTTTGAGAAAGT 1327
Db |||||
1654 TATTTATATATATATATTTAATATACAAATTTTATATATATATATATATATATATATATATA 1595
Qy 1328 TTTGAAATGTATAGAAATGAGTGGGACATCCATAAAAGGAAAGTGTATAGAAATTAATATGG 1387
Db |||||
1594 TTTATATATATATATATTTGATGATCATACATAATTTATATATATATATATATATATATATAT 1535
Qy 1388 GACAGAGGAGTAATACCTTTATGATATATAAAATTTTGTATTTGATTTTGCATTAAGATT 1447
Db |||||
1534 TTAATATATGAACTATTAATAAGA 1475
Qy 1448 ATAAATCTATGTTATATATGATA-ATATAATTTTAAAAATAATATACATATATATATCTGATT 1506
Db |||||
1474 ATTATATATATATGTTAT 1415

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:35:24 ; Search time 7744.85 seconds
(without alignments)
9654.718 Million cell updates/sec

Title: US-09-806-197-1
Perfect score: 2052
Sequence: 1 catggtgcccacacagcaca.....gcaataacattcttaaatc 2052

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : .EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| C 1 | 97.8 | 4.8 | 1101 | CNS00EVL | AL069706 Drosophil |
| C 2 | 97.2 | 4.7 | 1101 | CNS0039G | AG033921 Drosophil |
| C 3 | 95.6 | 4.7 | 1542 | AG386981 | Mus muscu |
| C 4 | 95 | 4.6 | 1101 | CNS0039G | AL063921 Drosophil |
| C 5 | 94.6 | 4.6 | 1758 | CL509408 | SAIL 811 |
| C 6 | 93.6 | 4.6 | 1896 | CG753083 | CG753083 P048-1-CO |
| C 7 | 93 | 4.5 | 1202 | CC262481 | CH261-167 |
| C 8 | 92.4 | 4.5 | 1101 | CNS0021J | AL061936 Drosophil |
| C 9 | 90.8 | 4.4 | 1101 | CNS00EVL | AL069706 Drosophil |
| C 10 | 90.4 | 4.4 | 1608 | CL118721 | ISB1-72J8 |
| C 11 | 88.8 | 4.3 | 1067 | CD386564 | AGENCOURT |
| C 12 | 88.8 | 4.3 | 1268 | AG347098 | Mus muscu |
| C 13 | 88.8 | 4.3 | 1392 | CG757503 | P052-4-CO |
| C 14 | 88 | 4.3 | 1277 | CC253231 | CH261-180 |
| C 15 | 88 | 4.3 | 1696 | AG346840 | Mus muscu |
| C 16 | 87.6 | 4.3 | 1275 | CL033318 | CH216-36F |
| C 17 | 87.2 | 4.2 | 887 | AG526041 | Mus muscu |
| C 18 | 87 | 4.2 | 1275 | CL033318 | CH216-36F |
| C 19 | 87 | 4.2 | 2087 | AG333887 | Mus muscu |
| C 20 | 86.8 | 4.2 | 1348 | CG749499 | P043-4-AO |
| C 21 | 85.6 | 4.2 | 1350 | CL019486 | CH216-5G1 |
| C 22 | 85.4 | 4.2 | 1780 | AG320553 | Mus muscu |
| C 23 | 85.2 | 4.2 | 1320 | CL103881 | ISB1-42C8 |
| C 24 | 85 | 4.1 | 1539 | AG340947 | Mus muscu |

| | | | | | | |
|------|------|-----|------|---|----------|--------------------|
| C 25 | 84.8 | 4.1 | 1254 | 9 | AG349719 | Mus muscu |
| C 26 | 84.6 | 4.1 | 987 | 9 | CNS014PQ | Drosophil |
| C 27 | 84.6 | 4.1 | 1364 | 9 | CG757970 | P053-2-CO |
| C 28 | 84 | 4.1 | 1101 | 9 | CNS017KE | Drosophil |
| C 29 | 83.8 | 4.1 | 822 | 7 | CK416977 | AUF-1pinc |
| C 30 | 83.6 | 4.1 | 1489 | 9 | AG350139 | AG350139 Mus muscu |
| C 31 | 83.4 | 4.1 | 1101 | 9 | CNS00PMC | AL070972 Drosophil |
| C 32 | 83 | 4.0 | 1211 | 9 | AG349657 | Mus muscu |
| C 33 | 83 | 4.0 | 1592 | 9 | CG750135 | P044-3-DO |
| C 34 | 82.8 | 4.0 | 1210 | 9 | CG749728 | P044-1-CO |
| C 35 | 82.6 | 4.0 | 1506 | 9 | AG278469 | Mus muscu |
| C 36 | 81.6 | 4.0 | 1092 | 9 | CNS020K7 | AL175636 Tetraodon |
| C 37 | 81.6 | 4.0 | 1101 | 9 | CNS00807 | AL069440 Drosophil |
| C 38 | 81.6 | 4.0 | 1355 | 9 | AG346348 | Mus muscu |
| C 39 | 81.6 | 4.0 | 1745 | 9 | AG338221 | Mus muscu |
| C 40 | 81.4 | 4.0 | 812 | 8 | BH178455 | O11_J_02- |
| C 41 | 81.4 | 4.0 | 812 | 9 | CNS07KRM | AL615412 T3_end of |
| C 42 | 81.4 | 4.0 | 1050 | 9 | CNS013NS | AL103090 Drosophil |
| C 43 | 81.4 | 4.0 | 1391 | 9 | CG754863 | P050-2-GO |
| C 44 | 81.2 | 4.0 | 1094 | 9 | CNS012FZ | AL101513 Drosophil |
| C 45 | 81.2 | 4.0 | 1715 | 9 | AG288305 | Mus muscu |

ALIGNMENTS

RESULT 1
CNS00EVL/c

LOCUS

DEFINITION

CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC:
BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL069706.1 GI:4949849
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosagawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source

1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR29B23"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN

Query Match 4.8%; Score 97.8; DB 9; Length 1101;
Best Local Similarity 34.8%; Pred. No. 8.4e-08;

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Hattori.M., Toyoda.A., Noguchi.H., Kojima.T. and Sakaki.Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1542)
Hattori.M., Toyoda.A., Noguchi.H., Kojima.T. and Sakaki.Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan 305-0856, Japan
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : TJ

LIBRARY : pBACs3.6

Vector : EcoRI

R.Site 1 : EcoRI.

R.Site 2 : EcoRI.

Location/Qualifiers

1. .1542

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMg01-201G10.TJ"

/sex="male"

/tissue_type="mixture of kidney and spleen"

/clone_lib="MSMg01 Mouse Male BAC Library"

FEATURES

source

ORIGIN

Query Match 4.7%; Score 95.6; DB 9; Length 1542;
Best Local Similarity 44.3%; Pred. No. 2.1e-07;
Matches 420; Conservative 0; Mismatches 523; Indels 6; Gaps 1;
QY 750 ATAAATAGTAGGATCTACATGACATTTAAATAGAGCTATTAATTAAGTTACTAATAA 809
DB 1103 ATAAGAAATTAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 1044
QY 810 ATAAGAGAGGTTAGTAAACAGAGAGGCTTAAACAGAGCTGCTGCTGCTTTAG 869
DB 1043 AATAATANAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 984
QY 870 TTGTTGTGAGCTCATTTCTTTAAAGTAATGTAAAGCTGATCTAAAGCACATAGAAATTTA 929
DB 983 TATAATAATATATATATATATATTAATAATAATAATAATAATAATAATAATAATAATA 924
QY 930 GTACAGGTTAAACTTTTACAGAAATTTATTAACGAAATCATTTTATACATGCTCT 989
DB 923 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 864
QY 990 CTCGGCTGCTATTAATAGGATCACTTACTGATCATCATTAATTAATTAATTAATTA 1043
DB 863 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 804
QY 1044 AAAACAAATTCATAGATAGATAAATAATCTTACATGAAAGAGACAAATGCTCTTTGAA 1103
DB 803 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 744
QY 1104 AAAACAAATAGGTACTCCCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1163
DB 743 TAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 684
QY 1164 ACTAAGAAATCTATAGATATGTAGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
DB 683 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 624
QY 1224 GTAGCGGGACCCCAATATATAATTTGATAGATTTAGAGAGAGAGAGAGAGAGAGAG 1283

DB 623 ATAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 564
QY 1284 GGGTGGGATTTTAT 1343
DB 563 TATATATAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 504
QY 1344 TTGAGTGGGACATCCATAAAGGAGAGTGTATAGAAATTAATTAATTAATTAATTA 1403
DB 503 TAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 444
QY 1404 CCTTATGAT 1463
DB 443 ATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 384
QY 1464 ATGAT 1523
DB 383 ATAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
QY 1524 TATAATTTTACAACTAGTATATATATATATATATATATATATATATATATATAT 1583
DB 323 AT 264
QY 1584 TTTGTAAACACAGCTTCGGTCAAATGGGAGTTCATGCTATTCATATATATATAT 1643
DB 263 AATAATTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 204
QY 1644 AGTAAATTTTAAATTAATTTGTTATTTTGTTCAGAAATTTTAAATAA 1692
DB 203 AAAAT 155

RESULT 4

CNS0039G

LOCUS

DEFINITION

CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazuo Oosawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain v2; cn bw ap, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

Location/Qualifiers

1. .1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR08K10"

FEATURES

source


```
Db 1413 TAAATTTTATTTTAAATATTTTAAATATATTTTAAATTTNANNAAATATAAATATATAT 1354
Qy 1010 GGATCAGCTTACTGATCATCCATTAACCTTCTTAAACCAATTCATGAGATATAATAT 1069
Db 1353 AAAAATATNATAATTTAAATATTTAAATATTTTAAATTTTAAATTTTAAATTTAT 1294
Qy 1070 CTTACATGAAAGAGGAGCAATGCTCTTTGAAAAAACAATAGGTACTCCCTCCGTC 1129
Db 1293 AAAAAATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1234
Qy 1130 CTCTCAATATGATACATATGATGAGTTCGACGAGACTAAGAAAAATGATATAAGTAAGT 1189
Db 1233 AAAAATATATATATATATATATATATATATAAATAAATAAATAAATAAATAAAT 1174
Qy 1190 AGAGTAAAGAAAGAGAGAAAGTGGTAAAGTAGCGGACCCACCAATATATAAT 1249
Db 1173 ATAAATATAATATAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1123
Qy 1250 GATGATTTAGAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 1309
Db 1122 TATAAATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1063
Qy 1310 TTACTATTTTGAGAAAGTTTGAAGTGTATAGATTTGAGTGGGACATCCATAAAGGAA 1369
Db 1062 TTATATAATATTAATATATATTTTATTTTATATATATATAAATAAATAAATAAAT 1003
Qy 1370 GTGTATAGAAATTAATGGGACAGAGGAGTATACCTTTTATGATATATAAATTTTGT 1429
Db 1002 TTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 943
Qy 1430 TTTTGATTTCAAGATTAATAAATCTA--TGTATAAATGATAAATAAATAAATAAATA 1487
Db 942 AATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 883
Qy 1488 TACTATATTT--AATCTGATTTAGTCGATTTACCGCTTTTATAATTTTACATCTAGT 1544
Db 882 AATAAATTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 823
Qy 1545 AATATGAATAATCAGTTATCTGAAAAAGCAATAATATCTTTTGTAAAAACAGCGTTC 1604
Db 822 ATNATNAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 763
Qy 1605 AAATGGGAGTTCATGTGATTTCAATAGTTTAAATATAAAGTAAATTTTAAATTAAT 1664
Db 762 ATTATATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 703
Qy 1665 TTATTTTGTTCAGAAATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1704
Db 702 ATATATAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 663
```

```
RESULT 7
LOCUS CC262481 1202 bp DNA linear GSS 13-MAY-2003
DEFINITION CH261-167M9_Sp6.1 CH261 Gallus gallus genomic clone CH261-167M9,
genomic survey sequence.
ACCESSION CC262481
VERSION CC262481.1 GI:30607397
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 1202)
Krenitzki, C., Higginbotham, J., Wylie, K., Carter, J., McPherson, J.,
Warren, W., Graves, T., Hards, E. and Wilson, R.
Gallus gallus BAC End Reads
Unpublished (2003)
Contact: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
```

```
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: Sp6 ATTAGGTGACACTATAG
Class: BAC ends
High quality sequence start: 30
High quality sequence stop: 105.
FEATURES
source
1..1202
location/Qualifiers
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strain="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-167M9"
/sex="Female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/notes="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: http://www.chori.org/bacpac"
```

ORIGIN

```
Query Match 4.5%; Score 93; DB 8; Length 1202;
Best Local Similarity 47.3%; Pred. No. 6.3e-07;
Matches 393; Conservative 0; Mismatches 420; Indels 17; Gaps 4;
Qy 886 TCCTTAAAGTAATGTAAACCTGATCTAAAGCACATAGAAATTTAGTACAGGTAAACCTT 945
Db 1174 TATNAAAAAATAATTAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1115
Qy 946 TTACAAGATTTATATAACGAAATCATTTTATACATCTCTCGGCTGCTATTATA 1005
Db 1114 ATAAAAAATTTTATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1055
Qy 1006 ATAGGGATCACTTCTGATCATCCATTAACCTTTGTTAAAAACAATTTCAATGAGATAA 1065
Db 1054 AATAATAATAATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 1003
Qy 1066 ATATCTTCAATGAAAGAGGACATGCTCTTTGAAAAACAATAAGGTACTCCCTCC 1125
Db 1002 ATATNTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 943
Qy 1126 GTCCCTCTGAATCTATACATATGATTTGGACACGAGACTAAGAAAAATGATATAAGTA 1185
Db 942 ATATTTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 883
Qy 1186 ATGTAGAGTAAAAAGAGAGAAAGTGGTAAAGTAGCGGACCCACCATATATAT 1245
Db 882 ATTAATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 823
Qy 1246 AATTGATAGATTAGAAAAAGTAGTTGAAAGTAGTGGGTGGGTGGGTGGGTGGGTGGGT 1305
Db 822 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 763
Qy 1306 AAATTTTACTATTTTGGAGAAAGTTTGAATCTATAGAAATTCAGTGGGACATCCATAAAG 1365
Db 762 TAATAAATTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 703
Qy 1366 GAAAGTGATAGAATTTAAATGGGACAGAGGAGTAATACCTTTATGATATATAAATTTT 1425
Db 702 AAAAA--TAAAAATNTAATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATA 645
Qy 1426 GTTATTTTGATTTTCATAAGATTTAAATCTATGTTTATGATATAAATAAATAAATAAAT 1485
Db 644 TATTTATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 585
Qy 1486 AATACTATTAATTTCTGATTAGTTCGATTCACCGCTTTTATAATTTTACAATACAGTA 1545
Db 584 AAAAAAATTTTATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 531
Qy 1546 ATATGAATAAATCAGTTATCTGAAAGCAATAAATATCTTTGTAAAAACAGCGTTCGGTCA 1605
Db 530 ATATATAAATTTTATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 471
```

```

QY 1606 AATGGGAAGTTCATGTGTTTCAATAGT-TTTAATATAAAAGTAATTTTAAATTAATG 1664
Db 470 AAAAAATAATATATATTTTAAATTTTAAATTTTAAATTAATAAATAATTT 411
QY 1665 TTATTTTGTTCAGAAATTTAAATTAATTTAGCATGGGAATTTCA 1714
Db 410 TTTTATATAAATTAATTAATATATATATAATAAATAATTTATTAATAATTA 361

RESULT 8
CNS0021J 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL061936.1 GI:4940214
VERSION GSS.
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila\_bac.htm.

FEATURES
source Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_lib="BACR05N11"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 4.5%; Score 92.4; DB 9; Length 1101;
Best Local Similarity 39.4%; Pred. No. 8.2e-07;
Matches 349; Conservative 65; Mismatches 469; Indels 2; Gaps 2;

QY 868 AGTTGTTGAGCTCATTTCTTTAAAGTAATGTAACGTATCTAAAGCACATAGAAATT 927
Db 138 AGTGTTCATGAAAGTTTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 197

QY 928 TAGTACAGGTAAACCTTTTACAGAAATTTATATTAACGAAATCATTTTATACATCT 987
Db 198 NAAAAAANAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 257

QY 988 CTCGCGCTGTCATTATAATAGGATCACTTACTGATCATCTCAATTAACCTTTGTTAAA 1047
Db 258 AAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 317

QY 1048 CAAATTCATGAGATTAATATCTTCAATGAAAGACGACATGCTCTTTTGAAGAAA 1107
Db 318 CAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 377

```

```

QY 1108 CAAATAGGTACTCCCTCCGTCCTGAAATCTATACATATGATGATGGACGGAGCTA 1167
Db 378 AAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 437
QY 1168 AGAAAAATGTATAAGTAAATGATAGTAAAAAGAAAGAAAGAAAGTGGGTAAAGTAG 1227
Db 438 AAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 497
QY 1228 CGGGACCCACCAATATATATGATAGATTTTAAAGAGTAGTTGAAAGTAGTGGGTGG 1287
Db 498 AAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 557
QY 1288 GGGATTTTATATATAAATTTTACTATTTTTCAGAAAGTTTTGAATGCTATAGAAATGA 1347
Db 558 TWTATATTTTAAATTTTAAAWAAATTTTAAATAAAAWATTTTAAWTTTTAAATTA 617
QY 1348 GTGGGACATCCATAAAGAAAGTGTATAGATTAATTAATGGGACAGAGGAGTAATACCTT 1407
Db 618 AAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 677
QY 1408 TATGATATAT-AAATTTTGTATTTTGTATTTTTCATAAGATTAATATCTATGTTAAAG 1466
Db 678 TWTATTAATTAATTTTAAAWAAATTTTAAATAAAAWATTTTAAWTTTTAAATTA 737
QY 1467 ATAATATATATTTTAAAAAATACTATATTAATTTCTGATAGTTCGATTAACCGCTTTTAT 1526
Db 738 AAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 797
QY 1527 AATTTTCAATCTAGTAATATGAATTAATCAAGTATCTGAAAGCAATATATCTTT 1586
Db 798 TTTAATTTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGTAAAAAAGT 857
QY 1587 GTAAACAGCGTCCGTCAGAAAGTTCATGTTTCAATAGTATTTTATATATAAAG 1646
Db 858 TTTAATTTTAAATTTTAAAWAAATTTTAAATAAAAWATTTTAAWTTTTAAATTA 916
QY 1647 TAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTT 1706
Db 917 AATATGKAGATGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGKKGK 976
QY 1707 GAAGTTTCAGGGCATCATTTGACGACGACGACGACGACGACGACGACGACGACGAC 1751
Db 977 TAAAAAADKAKGAAAAATTAAGAAAGKATGAAAAAGTGKDAATATT 1021

```

```

RESULT 9
CNS00EVL 1101 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC:
DEFINITION BACR29B23 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL069706.1 GI:4949849
VERSION AL069706
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1101)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Oosogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of

```


1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: hattori@sec.riken.jp, URL: http://hgp.gsc.riken.go.jp/,
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT
Clones are derived from the mouse BAC library MSMG01. For BAC
library availability, please contact Kunia Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp

PRIMERS

Sequencing : T7

LIBRARY

Vector : pBac3.6

R.Site 1 : EcoRI

R.Site 2 : EcoRI

Location/Qualifiers

1. .1268

/organism="Mus musculus molossinus"

/mol_type="genomic DNA"

/sub_species="molossinus"

/db_xref="taxon:57486"

/clone="MSMG01-142102.T7"

/sex="male"

/tissue type="mixture of kidney and spleen"

/clone_lib="MSMG01 Mouse Male BAC Library"

ORIGIN

Query Match 4.3%; Score 88.8; DB 9; Length 1268;
Best Local Similarity 45.5%; Pred. No. 3.7e-06;
Matches 471, Conservative 0; Mismatches 551; Indels 14; Gaps 4;

```

QY 1030 ATTAAACCTGTTAAACAAATCAATGAGATAAAATATCTTCAACATGAAAGAGAC 1089
    |||||
DB 1215 ATTAAATATNATAATAAAATATATATAATATAATATAATATAATATAATATA 1156
    |||||
QY 1090 ATGTCTCTTGAAGAAACCAATAGTACTCCCTCGCTCCTGAAAT--GTATACAT 1146
    |||||
DB 1155 AATAATATATATAATATAAAATATAATATAATATAATATAATATAATATAAT 1096
    |||||
QY 1147 ATGGATTGGACACGGAGACTAGAAAAATGTATAAGTAAGTAGAGTAAAAAGAGAG 1206
    |||||
DB 1095 ATTAAATATTATTATAATATAATATAATATAATATAATATAATATAATATAAT 1036
    |||||
QY 1207 AAGAAAGAGTGGTAAAGTAGCGGACCCACCAATATATATATATAGTAGATTAGAAAAGT 1266
    |||||
DB 1035 AATAAAATATAATATAATATAAAATATAAAATATAAAATATAATATAATATAAAAT 976
    |||||
QY 1267 AGTGAAGTAGTGGTGGGATTTTATATTATAAAATTTTACTATTTCGAGAAAG 1326
    |||||
DB 975 ATTATATAAATAATATTATAAATAATATTATAATATAAATAATATAAATAATATA 916
    |||||
QY 1327 TTTTGAATGTAGAAATGAGTGGGACATCCATAAAGGAAAGTGTATAGAAATTAATG 1386
    |||||
DB 915 ATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 856
    |||||
QY 1387 GGACAGAGGAGTAATACCTTTATGATATATAAATTTTGGTATTTTGAATTCATAAGAT 1446
    |||||
DB 855 TTAATAATAAATAAATAATTTTATATATAAATAAATAAATAAATAAATAAATA 796
    |||||
QY 1447 TATAATCTGCTTATATGATAAATAATTTTAAATAAATACTATATTAATCTGATT 1506
    |||||
DB 795 ATAAATAATATTAAATATAATATTATATAAATAAATAATTTTAAATAAATAAATA 736
    |||||
QY 1507 AGTCGATTACCGCCTTTTATATAATTTTCAATACTGAGTAATATGAATAAATCAGTTATCT 1566
    |||||
DB 735 AATAAATAAATTTATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 676
    |||||
QY 1567 GAAAGAGCAAT---AATATCTTTGTAAGAACACGCGTTCGGTCAATGGGAAGTTCAATGTT 1623
    |||||
DB 675 AAAATAAATAAATAAATAATATAATATAAATAAATAAATAAATAAATAAATA 616
    |||||
QY 1624 ATTCAATAGTTTTTAATAAATAAAGTAAATTTTAAATTAATTTGTTATTTTGTTCAGAAAT 1683
    |||||

```

```

Db 615 ATATTATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 556
    |||||
QY 1684 TTAATAAATAAATATTAGCATGGGAAGTTTCACGGGCATCAATTGA---GCAGCACTAGACT 1740
    |||||
Db 555 ATAAATATATATATATATATATATAAATAAATAAATAAATAAATAAATAAATAA 496
    |||||
QY 1741 GTTTGAACAATGATATGTCGGGTGACATCTATGACCTTTTCAACTCAAACTAGTAGTAATAAT 1800
    |||||
Db 495 ATATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 436
    |||||
QY 1801 GCATTCCTAGATACATCTTTTCAAAATTTCAACAAACACAGCTTTTAACTTTTCTT---T 1855
    |||||
Db 435 AGTAAATAAATAAATTAGGTATTTAAATAAATAAATAAATAAATAAATAAATAA 376
    |||||
QY 1856 CAACGGATTGGAATCCTTTTCTAAACTTTTTTAAATAAATAAATAAATAAATAA 1915
    |||||
Db 375 AAAAAATTTATAATATTTTATAAATAAATAAATAAATAAATAAATAAATAAATAA 316
    |||||
QY 1916 TTATCAACACCTCAACATTTGATGTTAGGTACTATATAATAGGTGCTCTTGGTGTCTTACT 1975
    |||||
Db 315 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 256
    |||||
QY 1976 ATCATCACATCAATCTTTACACCACAAACCTTGAGCTTAATTTTCTACTTATCTCAGCA 2035
    |||||
Db 255 TAATTATATTAAATTAATTAATATAAATAAATAAATAAATAAATAAATAAATAA 196
    |||||
QY 2036 ATAACTTTCTAAATAT 2051
    |||||
Db 195 ATAAATTAATATAT 180
    |||||

RESULT 13
CG757503/c 1392 bp DNA linear GSS 24-OCT-2003
LOCUS P052-4-C08.za Ppa EcoRI BAC Library Pristionchus pacificus genomic,
DEFINITION genomic survey sequence.
ACCESSION CG757503
VERSION CG757503.1 GI:37986131
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
          Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
          Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 1392)
AUTHORS Srinivasan,J., Sinz,W., Jesse,T., Wiggers-Pereboite,I., Jansen,K.,
          Buntjer,J., van der Meulen,M. and Sommer,R.J.
          An integrated physical and genetic map of the nematode Pristionchus
          pacificus
JOURNAL Mol. Genet. Genomics 269 (5), 715-722 (2003)
MEDLINE 22835951
PUBMED 12884007
COMMENT Contact: Sommer RJ
          Evolutionary Biology
          Max-Planck-Institute for Developmental Biology
          Spemannstr. 37-39, Tuebingen D-72076, Germany
          Tel: 00497071601371
          Fax: 00497071601498
          Email: ralf.sommer@uebingen.mpg.de
          Class: BAC ends.
FEATURES             Location/Qualifiers
     source            1..1392
                        /organism="Pristionchus pacificus"
                        /mol_type="genomic DNA"
                        /strain="California"
                        /db_xref="taxon:54126"
                        /clone_lib="Ppa EcoRI BAC Library"
                        /note="The library was generated by a partial digest of
                        the genomic DNA with EcoRI and cloning into the BAC
                        vector."
ORIGIN
Query Match 4.3%; Score 88.8; DB 9; Length 1392;

```

Best Local Similarity 43.8%; Pred. No. 3.7e-06;
Matches 543; Conservative 0; Mismatches 672; Indels 24; Gaps 3;

QY 481 ATGCTCAGCCATCAAAATTCACAAAACCCGACACACACTCTATCCACGCTACTATCTT 540
Db 1359 ATAATAAATATATAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1300

QY 541 TGGCCGAATGCTTCTCAAAATGTTTTTATATATGTAATAATGCCCATCCAGGATAAGT 600
Db 1299 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1240

QY 601 AAAATTCCTGTTTAAACGTTTGTATAATATATGTTTACACTTAACAAGGAGATATCGT 660
Db 1239 ATTATTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1180

QY 661 AATACTTTTAGACGACAGAGACTTAGTCAAAATGGAGCTGGTAAACAGCCTAGACT 720
Db 1179 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1120

QY 721 TGGTCACTGATAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGACATAAA 780
Db 1119 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1060

QY 781 ATTAGACTTATTAATTAAGTTACTAATAATAAGAGGTTAGTAAACGAAAGCAGGTA 840
Db 1059 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1000

QY 841 AAAACAGAGCTGCTGCTGCTGTTTGTAGTGTGTTGAGCTCATCTTTTAAAGTAAG 900
Db 999 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 940

QY 901 TAAACTGATCTAAGCACATAGA-----AATTTAGTACAGGTTAAACTTTTAC 949
Db 939 AATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 880

QY 950 AAGAATTTATTAACGAAATCATTTTATAACATGTCTCTCGGCTGCTATTATATAG 1009
Db 879 AATAAATATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 820

QY 1010 GGATCACTACTCATCCATTAACACCTTGTAAACCAATTCATGAGATAAATAAT 1069
Db 819 AATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 760

QY 1070 CTTTCAATGAAAGAGGACATGTCTCTTTGAAAAAACAATAGGTACTCCCTCGCTCC 1129
Db 759 ATTATTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 700

QY 1130 CTCTGAAATGTATACATATGGAATGGACCGGAGACTAAGAAAAATGTATAAAGTAATGT 1189
Db 699 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 640

QY 1190 AGAGTAAAAAGAAAGAAAGAAAGTGGTAAAGTAGCGGACCCACCACCAATATAAT 1249
Db 639 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 580

QY 1250 GATAGATTTAGAAAAAGTAGTTGAAA-----GTAGTGGTGGGTGGGATTTTATATATA 1305
Db 579 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 520

QY 1306 AAATTTACTATTTTGAGAAAGTTTTGAAATGTATAGAAATGAGTGGGACATCCATAAAG 1365
Db 519 ATTATATATATTATATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 460

QY 1366 GAAAGTGTATAGAAATTAATGGACAGAGGAGTAATACCTTTATGATATATAAATTTT 1425
Db 459 TAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 400

QY 1426 GTTATTTTGAATCAAGATTATAATCTATCTGTTATAATGATAATATAATTTTAAATAAT 1485
Db 399 AATAATTTATTTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 340

QY 1486 AATPACTATATTAATCTGATTAGTCGATTCACGCTTTTATATATTTTACAACTAGTA 1545

Db 339 AAATATATAAATTAATTTATATATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 280

QY 1546 ATATGAATAAATCAGTTATCTGAAAGCAAAATAATATCTTTGTAAACACGCTTCGGTCA 1605

Db 279 TAAAAATTTATATAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 220

QY 1606 AATGGGAAGTTCATGTGTTTCAATAGTTTAAATATAAAGTAAATTTTAAATTTAATTTGT 1665

Db 219 AAATATATATTAATTTT-----TAATAAATTTAAATAAATTTAAATAATATA 169

QY 1666 TAATTTTGTTCAGAAATTTAAATAAATAATTTATGAGCAT 1704

Db 168 TATTATTATATAAATAAATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 130

RESULT 14
LOCUS CC253231 1277 bp DNA linear GSS 13-MAY-2003
DEFINITION CH261-180N11 RM1.1 CH261 Gallus gallus genomic clone CH261-180N11,
genomic survey sequence.
ACCESSION CC253231
VERSION CC253231.1 GI:30589981
KEYWORDS GSS.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1277)
AUTHORS Kremitzki,C., Higginbotham,J., Wylie,K., Carter,J., McPherson,J.,
Warren,W., Graves,T., Mardis,E. and Wilson,R.
Gallus gallus BAC End Reads
Unpublished (2003)
CONTACT: Richard K. Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 182000 Std Error: 0.00
Seq primer: RM1 TAGCACTCACTATAGGAGA
Class: BAC ends
High quality sequence start: 24
High quality sequence stop: 85.

FEATURES
source
1..1277
/organism="Gallus gallus"
/mol_type="genomic DNA"
/strains="Red Jungle Fowl"
/db_xref="taxon:9031"
/clone="CH261-180N11"
/sex="female"
/cell_line="UCD001, inbred 256"
/clone_lib="CH261"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CH261 Female Chicken library - for library and clone
ordering information: <http://www.chori.org/bacpac>

ORIGIN
Query Match 4.3%; Score 88; DB 8; Length 1277;
Best Local Similarity 43.9%; Pred. No. 5.2e-06;
Matches 475; Conservative 0; Mismatches 600; Indels 6; Gaps 3;

QY 619 GTTGTGTTAATATATATGTTTACACTTACAAGAGGATATTCGTAATACTTTTAGACACAA 678

Db 41 GT 100

QY 679 GAGACTTAGGTCAAAATAGGACGCTGGTAAACAGCCTAGACTTGGTCACTGATAAATAGA 738

Db 101 GAACCTTTAAAAAANNCCAAAAAATACATATTTTATAAATAATATTTAAATAATATA 160

QY 739 TAATTTGTTAGTATAATATAGTAGGATCTCAATGACATTAATAAATAGACTTAAATTA 798

Db 161 TATTTTAAACAATTTTTTTTTGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAT 220

```
QY 799 GTTACTAATAAATAGAGAGTTAGTAAACAGAAAGCAGGTAAACACAGAGCTTGCTGC 858
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
221 ATTAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 280
QY 859 TGTGTGTTTGTGTTGTTGAGCTCATTTCTTTAAAGTAATGTAACATGATCTAAAGCAC 918
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
281 ATAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 340
QY 919 ATAGAAATTTAGTACAGGTTAAACCTTTTCAAGAAATTTATATTAACGAAATCATTTT 978
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
341 TATAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 400
QY 979 ATAAACATGCTCTCGGCTCATTAATAAGGATCAGTCTGATCATCCATTAACACC 1038
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
401 ATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 460
QY 1039 TTGTTAAAAACAATTCATAGAGATAAATCTCTCAATGAAGAAAGAGACATGCTCT 1098
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
461 TTTATATATTTTATTATTAATTAATAATTAATAATTAATAATTAATAATTAATAATTA 520
QY 1099 TTGAAAAACAATAGGTACTCCCTCCGTCCTCTGAAATGTATATATATGATGGAAC 1158
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
521 ATATTAACATATATATTTTAATAATTTAAATATATATATTAATAAATAAATAATATATA 580
QY 1159 CGGA-GACTAAGAAATATATAAGTATAGTAGTAAAGAAAGAGAAAGAAAGTG 1217
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
581 TAAATTTTAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAAT 640
QY 1218 GGTAAAGTAGCGGACCCACCAATATATAATTTGATAGATTTTGAAGAAAGTAGTGAAGTA 1277
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
641 ATAATAATAAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATAATTAATA 700
QY 1278 GTGGGTGGGTGGATTTTATATATTAATAAATTTACTATTTTGAAGAAAGTTTGAAGT 1337
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
701 ATTAATTTTAAATTTTATATAATTAATTAATAATTAATTTTAAATTTTAAATTTTAAAT 760
QY 1338 ATAGAATTCAGTGGGACATCCATAAAGGAAGTGTATAGATTAATTAAGTAAATTAAT 1397
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
761 AATATATATAATTTTAAATAAATTAATAATTAATAATTAATAATTAATAATTAATAAT 820
QY 1398 GTAATACCTTTATGATATATAAATTTTGTATTTTGTATTTTCAAGATTTATAAATCTAT 1457
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
821 AAAATATATAAATAATTAATAATTTTATATATAAATAAATAAATAAATAAATAATATAT 880
QY 1458 GTTATAATGATTAATAATTTTAAATAAATACTATTAATTCGTATGATGCGATTACC 1517
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
881 TATAAATTTAACATAAATATATATAATTTAAATTTATTAATTTAAATTTAAATTTAAAT 940
QY 1518 GCCTTTTATAATTTTACATACTGAGTATATGAATTAATCAGTTATCTGAAAGCAAT 1577
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
941 ATATTTTATA---TAATAAATAATTAATTTTATATAAATAAATAAATAAATAAATAAAT 997
QY 1578 AATATCTTTGTAAAAACAGCGTTTCGGTCAAATGGGAAGTTTCATGTGTTTCAATAGTTTA 1637
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
998 AAATTTTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1055
QY 1638 ATATAAAGTAATTTTAAATTAATTTTGTATTTTGTTCAGAAATTTAAATAAATATAT 1697
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1056 ATAAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1115
QY 1698 T 1698
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1116 T 1116
```

```
RESULT 15
AG346840/c
LOCUS Mus musculus molossinus DNA, clone:MSMg01-142C12.17, genomic survey
DEFINITION sequence.
ACCESSION AG346840
VERSION AG346840.1 GI:47920150
KEYWORDS GSS.
```

```
SOURCE
ORGANISM Mus musculus molossinus
Mus musculus molossinus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
BAC end Sequences of Library MSMg01
Unpublished
2 (bases 1 to 1696)
Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
Direct Submission
Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@psc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : T7
LIBRARY
Vector : pBACE3.6
R.Site 1 : EcoRI.
R.Site 2 : EcoRI.
FEATURES
Location/Qualifiers
source
1..1696
/organism="Mus musculus molossinus"
/mol_type="genomic DNA"
/sub_species="molossinus"
/db_xref="taxon:57496"
/clone="MSMg01-142C12.17"
/sex="male"
/tissue_type="mixture of kidney and spleen"
/clone_lib="MSMg01 Mouse Male BAC Library"
```

ORIGIN

```
Query Match 4.3%; Score 88; DB 9; Length 1696;
Best Local Similarity 47.3%; Pred. No. 5.2e-06;
Matches 464; Conservative 0; Mismatches 501; Indels 16; Gaps 6;
QY 715 TAGACTTGGTCACGTAAATAGATAATTTGTTAGTATAATATAGTAGGATCTACAATGAC 774
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1290 TATAATTAATAATAATACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1231
QY 775 ATTAAATTAGAGCTNTTAATTAAGTTACTTAATAAATAAGAGAGGTTAGTAAACAGAAAG 834
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1230 ATAAAAATTAACAAAAACAACATAATATACATAAATAATAAATAAATAAATAAATAA 1171
QY 835 CAGGTAAAAACAAGAGCTTGCTGCTGTGTTTGTAGTTGTTGCTGAGCTCATTTCTTAA 894
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1170 AATAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1111
QY 895 GTAATGTAACCTGATTAAGAGCACATAGAAATTTAGTACAGGTTTAAATTTTACAAGAA 954
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1110 ATATAAATAATCAACATACATCAAAATTTATAAAAAAATAAATAAATAAATAAATAA 1052
QY 955 TTTATNTTAAACGAAATCATTTTATAACATGCTCTCTGGGCTGTCNTTATTAATAGGATC 1014
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1051 ATCATATTTAAATAATATATCTATCAATAATAAACNCAAAACATAATATATAATATCACC 992
QY 1015 ACTTACTGATCATCCATTAATAAACCTTGTAAAAAATAATCAATGAGATAAATAATCTTAC 1074
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
991 ATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAATTT-- 934
QY 1075 AATGAAAGAGGACAATGTCTCTTTGAAAAAATAAATAAGTACTCCCTCCGCTCCCTCTG 1134
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
933 -----AATAAACAATAATATATAATATATATATATTAATAAATAAACAATAAATAATA 880
```


THIS PAGE BLANK (USPTO)

| Result No. | Query | Score | Query | | | DB | ID | Description |
|------------|-------|-------|-------|--------|----------|-------------|-------------|-------------|
| | | | Match | Length | DB | | | |
| 1 | 851 | 100.0 | 851 | 6 | E40088 | E40088 | Plant promo | |
| 2 | 123.6 | 14.5 | 739 | 6 | E09663 | E09663 | cdNA encodi | |
| 3 | 123.6 | 14.5 | 739 | 6 | E40092 | E40092 | Plant promo | |
| 4 | 123.6 | 14.5 | 739 | 6 | 120016 | 120016 | Sequence 3 | |
| 5 | 123.6 | 14.5 | 739 | 8 | D89388 | D89388 | Daucus caro | |
| 6 | 122.6 | 14.4 | 774 | 8 | AB127961 | AB127961 | Daucus ca | |
| 7 | 90 | 10.6 | 727 | 8 | DCU47087 | DCU47087 | Daucus caro | |
| 8 | 88.6 | 10.4 | 2048 | 6 | E40089 | E40089 | Plant promo | |
| 9 | 88.6 | 10.4 | 2048 | 6 | E40090 | E40090 | Plant promo | |
| 10 | 88.6 | 10.4 | 2052 | 6 | E40087 | E40087 | Plant promo | |
| 11 | 88.6 | 10.4 | 2052 | 6 | E40093 | E40093 | Plant promo | |
| 12 | 88.6 | 10.4 | 2056 | 6 | E40091 | E40091 | Plant promo | |
| 13 | 77.8 | 9.1 | 7218 | 6 | I66494 | I66494 | Sequence 14 | |
| 14 | 76.4 | 9.0 | 8005 | 8 | DCA18706 | Y18706 | Daucus caro | |
| 15 | 74.8 | 8.8 | 2215 | 8 | DCU56392 | U56392 | Daucus caro | |
| 16 | 74.8 | 8.8 | 2215 | 8 | S83359 | IEP4=42.6-k | IEP4=42.6-k | |
| 17 | 73.2 | 8.6 | 4886 | 8 | DARGCHS2 | D16255 | Carrot gCHS | |
| 18 | 71 | 8.3 | 2042 | 6 | AR076817 | AR076817 | Sequence | |
| 19 | 71 | 8.3 | 2042 | 6 | E15125 | E15125 | Promoter. 7 | |

QY 61 CTGAAGGACTAGCGAACTGGAACCTGGTCTCAGACGAGGATCACCTACGAGCAAGTGA 120
Db 61 CTGAAGGACTAGCGAACTGGAACCTGGTCTCAGACGAGGATCACCTACGAGCAAGTGA 120
QY 121 AGAAATGGCAACTTATTTGATGACTTGTGATATTTATGAGAAATTTACAATTTTGCAG 180
Db 121 AGAAATGGCAACTTATTTGATGACTTGTGATATTTATGAGAAATTTACAATTTTGCAG 180
QY 181 AAAAGCCAACTCTGGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGAA 240
Db 181 AAAAGCCAACTCTGGGTTTCAGGTTCTGGAATTAATTAAGCTTTTACATCACTATGAA 240
QY 241 GTGATACCTATTCGAAGCGAGTGGAGGAGAACTTGTAGTACGTGAATGAGAAATTCAG 300
Db 241 GTGATACCTATTCGAAGCGAGTGGAGGAGAACTTGTAGTACGTGAATGAGAAATTCAG 300
QY 301 CAACACAGTACAAGAAAGCTCTGGAAGTAATTTGAGAGTATGCAATTAAGGAGAAATTCGG 360
Db 301 CAACACAGTACAAGAAAGCTCTGGAAGTAATTTGAGAGTATGCAATTAAGGAGAAATTCGG 360
QY 361 GGATACGTCAAAGTATTTTACGACACAAATAGAAAAGCGAAGGATAAAGTGCTCT 420
Db 361 GGATACGTCAAAGTATTTTACGACACAAATAGAAAAGCGAAGGATAAAGTGCTCT 420
QY 421 ATGAATGGTGAAGAACTGGACGGTTTAGGTTAGCTTAAAAAGTGACTTCTTACTTGA 480
Db 421 ATGAATGGTGAAGAACTGGACGGTTTAGGTTAGCTTAAAAAGTGACTTCTTACTTGA 480
QY 481 AGTAATGAAGTGGAGTGAAGTAAAGTAAATTAATTAAGTATTTAAAGTGTTT 540
Db 481 AGTAATGAAGTGGAGTGAAGTAAAGTAAATTAATTAAGTATTTAAAGTGTTT 540
QY 541 GGAAGAAAGTAAAGTGTGAAGAAAGTTAGTACATTTTCTACTTCCAACTTATTTCTCA 600
Db 541 GGAAGAAAGTAAAGTGTGAAGAAAGTTAGTACATTTTCTACTTCCAACTTATTTCTCA 600
QY 601 CGACTTCTTAAAGTACTTCTTACTTTTACACAAACGGTCAAGGAAAGTGGAAAGCA 660
Db 601 CGACTTCTTAAAGTACTTCTTACTTTTACACAAACGGTCAAGGAAAGTGGAAAGCA 660
QY 661 AAAGCTGGAGTTACTTCTTATAAGAAATGTTTATCTAAATGAGAAATGACAAACACAGAA 720
Db 661 AAAGCTGGAGTTACTTCTTATAAGAAATGTTTATCTAAATGAGAAATGACAAACACAGAA 720
QY 721 ATGAGAAATGAATATGATTTAGTGTAAATATAGTGTATTTTATTTAAAAAGATCGCAT 780
Db 721 ATGAGAAATGAATATGATTTAGTGTAAATATAGTGTATTTTATTTAAAAAGATCGCAT 780
QY 781 CATTACCGCCAGATGAAGTTATTCATCACAACTCACAAAGTACAAAGAAAAGTTG 840
Db 781 CATTACCGCCAGATGAAGTTATTCATCACAACTCACAAAGTACAAAGAAAAGTTG 840
QY 841 CAATTCTGTCA 851
Db 841 CAATTCTGTCA 851

RESULT 2
E09663/c
LOCUS E09663 739 bp RNA linear PAT 29-SEP-1997
DEFINITION cDNA encoding novel protein of carrot root.
ACCESSION E09663
VERSION E09663
KEYWORDS E09663.1 GI:22026290
SOURCE JP 1995188288-A/1.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 739)
AUTHORS Kubo,M. and Oita,K.
TITLE PROTEIN, SPECIFICALLY EXPRESSIBLE IN ROOT PART OF CARROT AND HAVING 16KD MOLECULAR WEIGHT, ITS GENE AND PLASMID CONTAINING THE SAME GENE
JOURNAL Patent: JP 1995188288-A 1 25-JUL-1995;

SUMITOMO CHEM CO LTD
OS Daucus carota L. (carrot)
PN JP 1995188288-A/1
PD 25-JUL-1995
PF 24-DEC-1993 JP 1993327943
PI KUBO MIYOSHI, OITA KENJI
PC C07K14/415,C12N15/09//A01H1/00,C12P21/02;
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC Feature is identified by experimental;
FH Key Location/Qualifiers
FT source 1..739
FT /organism='Daucus carota L.'
FT /strains='Kurodagoun'
FT /tissue_type='root'
FT /clone='PC16'
FT 3'UTR 479..717
FT polyA_site 718..739
FT CDS 14..478
FT /product='novel protein of carrot root'.
FEATURES
source 1..739
/organism='unidentified'
/mol_type='genomic RNA'
/db_xref='taxon:32644'
ORIGIN
Query Match 14.5%; Score 123.6; DB 6; Length 739;
Best Local Similarity 96.9%; Pred. No. le-13; Indels 0; Gaps 0;
Matches 126; Conservative 0; Mismatches 4;
QY 722 TGAGAAATGAATATGATTTAGTGTAAATAGTGTATTTTATTTAAAAAGATCGCATAC 781
Db 718 TGAGAAATGAATATGATTTAGTGTAAATAGTGTATTTTATTTAAAAAGATCGCATAC 659
QY 782 ATTACCGCCAGATGAAGTTATTCATCACAACTCACAAAGTACAAAGAAAAGTTGC 841
Db 658 ATTACCGCCAGATGAAGTTATTCATCACAACTCACAAAGTACAAAGAAAAGTTGC 599
QY 842 AATTCTGTCA 851
Db 598 AATTCTGTCA 589
RESULT 3
E40092/c
LOCUS E40092 739 bp DNA linear PAT 31-JAN-2002
DEFINITION Plant promoter and terminator.
ACCESSION E40092
VERSION E40092.1 GI:18627208
KEYWORDS JP 2000166577-A/6.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 739)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 6 20-JUN-2000;
SUMITOMO CHEM CO LTD
OS Daucus carota L.
PN JP 2000166577-A/6
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//C12N5/10,C12R1.91), PC C12N15/00,
PC C12N5/00,(C12N5/00,C12R1.91)
CC CC
FH Key Location/Qualifiers


```
FT source 1..739
FT /organism="Daucus carota L."
FEATURES
  source
    1..739
    /organism="unidentified"
    /mol_type="genomic DNA"
    /db_xref="taxon:32644"
ORIGIN
  Query Match 14.5%; Score 123.6; DB 6; Length 739;
  Best Local Similarity 96.9%; Pred. No. 1e-13;
  Matches 126; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 722 TGAGATGAATGATGATTGTTTAAATAGTGTATTTTAAAGAGATCGCATAC 781
Db 718 TGAGATGAATGATGATTGTTTAAATAGTGTATTTTAAAGAGATCGCATAC 659
QY 782 ATTACGAGCAGATGAAGTTATTCATCACAACATCACAAGTACAAAAGAAAAGTTGC 841
Db 658 ATTACGAGCAGATGAAGTTATTCATCACAAGTACAAAAGTACAAAAGAAAAGTTGC 599
QY 842 AATTCTGTCA 851
Db 598 AATTCTGTCA 589
RESULT 4
120016/c
LOCUS 120016 739 bp DNA linear PAT 07-OCT-1996
DEFINITION Sequence 3 from patent US 5512484.
ACCESSION 120016
VERSION 120016.1 GI:1600371
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 739)
AUTHORS Yamamoto,M. and Oheda,K.
TITLE Carrot 16 KD protein, gene coding for said protein and plasmid
  containing said gene
JOURNAL Patent: US 5512484-A 3 30-APR-1996;
FEATURES
  source
    1..739
    /organism="unknown"
    /mol_type="unassigned DNA"
ORIGIN
  Query Match 14.5%; Score 123.6; DB 6; Length 739;
  Best Local Similarity 96.9%; Pred. No. 1e-13;
  Matches 126; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 722 TGAGATGAATGATGATTGTTTAAATAGTGTATTTTAAAGAGATCGCATAC 781
Db 718 TGAGATGAATGATGATTGTTTAAATAGTGTATTTTAAAGAGATCGCATAC 659
QY 782 ATTACGAGCAGATGAAGTTATTCATCACAACATCACAAGTACAAAAGAAAAGTTGC 841
Db 658 ATTACGAGCAGATGAAGTTATTCATCACAAGTACAAAAGTACAAAAGAAAAGTTGC 599
QY 842 AATTCTGTCA 851
Db 598 AATTCTGTCA 589
RESULT 5
D88388/c
LOCUS D88388 739 bp mRNA linear PLN 07-FEB-1999
DEFINITION Daucus carota mRNA for cr16, complete cds.
ACCESSION D88388
VERSION D88388.1 GI:1663521
KEYWORDS cr16.
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Daucinae; Scandiceae; Daucinae; Daucus.

1 (bases 1 to 739)
Yamamoto,M., Torikai,S. and Oeda,K.
A major root protein of carrot with high homology to intracellular pathogen-related(PR) proteins and pollen allergens
Unpublished
2 (bases 1 to 739)
Torikai,S.
Direct Submission
Submitted (14-OCT-1996) Satomi Torikai, Sumitomo Chemical Co., Ltd., Biotechnology Laboratory; 4-2-1,Takatsukasa, Takarazuka, Hyogo 665, Japan (E-mail:torikai@bio.sumitomo-chem.co.jp, Tel:0797-74-2059, Fax:0797-74-2133)

FEATURES
source
1..739
/organism="Daucus carota"
/mol_type="mRNA"
/db_xref="taxon:4039"
/tisue_type="root"
gene
1..739
/gene="CR16"
CDS
14..478
/gene="CR16"
/note="A major root protein"
/codon_start=1
/product="cr16"
/protein_id="BAAL3604.1"
/db_xref="GI:1663522"
/translation="MGAQSHSLTSSVSAEKIFGIVLDVTVTPKAAPGAYKSVDV KDGCGAGTVRIITLPEGSPTSMVTVRTDAVNKEALTYDSTVDIGDILFEFISIEIHM VVPTADGSGSITKTATHTKGDVVPENIKFPADAQNTALFKAENYLIAN"

ORIGIN
Query Match 14.5%; Score 123.6; DB 8; Length 739;
Best Local Similarity 96.9%; Pred. No. 1e-13;
Matches 126; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 722 TGAGATGAATGATGATTGTTTAAATAGTGTATTTTAAAGAGATCGCATAC 781
Db 718 TGAGATGAATGATGATTGTTTAAATAGTGTATTTTAAAGAGATCGCATAC 659
QY 782 ATTACGAGCAGATGAAGTTATTCATCACAACATCACAAGTACAAAAGAAAAGTTGC 841
Db 658 ATTACGAGCAGATGAAGTTATTCATCACAAGTACAAAAGTACAAAAGAAAAGTTGC 599
QY 842 AATTCTGTCA 851
Db 598 AATTCTGTCA 589
RESULT 6
AB127961/c
LOCUS AB127961 774 bp mRNA linear PLN 26-MAY-2004
DEFINITION Daucus carota DcPRP mRNA for pathogenesis-related protein, complete cds.
ACCESSION AB127961
VERSION AB127961.1 GI:39104473
KEYWORDS
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Daucinae; Scandiceae; Daucinae; Daucus.

1
Sano,T., Nishimoto,M., Saburi,W., Kimura,A., Yasuda,H., Uchibatake,M., Ohwada,T. and Masuda,H.
Isolation and Characterization of cDNA Encoding P-19.5 Protein Accumulated Preferentially at Early Stage of Carrot Somatic Embryogenesis

JOURNAL Plant Sci. (2004) In press
REFERENCE 2 (bases 1 to 774)
AUTHORS Sano,T. and Masuda,H.
TITLE Direct Submission
JOURNAL Submitted (04-DEC-2003) Takuma Sano, Obihiro University of Agriculture and Veterinary Medicine, Laboratory of Biological Function; Nishi 2-11, Inada-Cho, Obihiro, Hokkaido 080-8555, Japan (E-mail:anos@octv.ne.jp, URL:http://www.obihiro.ac.jp/~kinoukaihatu/, Tel:81-155-49-5115(ex.5553), Fax:81-155-49-5577)
FEATURES Location/Qualifiers
source 1. .774
/organism="Daucus carota"
/mol_type="mRNA"
/db_xref="taxon:4039"
/note="pathogenesis-related protein accumulated in all stages of somatic embryogenesis and plantlet, identical to cr16"
gene 1. .774
/gene="DcPRP"
CDS 72. .536
/gene="DcPRP"
/codon_start=1
/product="pathogenesis-related protein"
/protein_id="BAD04049.1"
/db_xref="GI:39104474"
/translation="MGAQSHSLTSSVSASKIFSGIVLVDVTIPKAAPGAYKSDV KDGAGVRIITLPEGSPITSMVTVRDANKALTYDSTVIDGDLLEFTESIETHM VVVPADGGSITKTTAIFHTKGDVAVPEENIKFADAQNTALFKAEAYLIAN"
ORIGIN
Query Match 14.4%; Score 122.6; DB 8; Length 774;
Best Local Similarity 96.9%; Pred. No. 1.6e-13;
Matches 125; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 723 GAGATGAATGATGATTATTTGGTTTAAATAGTGTATTTATTTAAAAAGATCGCATACA 782
Db 774 GAGATGAATGATGATTATTTGGTTTAAATAGTGTATTTATTTAAAAAGATCGCATACA 715
Qy 783 TTACCAGCCAGATGAAGTTATTCATCACAACTCACAAAGTACAAAGAGAAAAAGTTGCA 842
Db 714 TTACCAGCCAGATGAAGTTATTCATCACAAAGTACAAAGTACAAAGAGAAAAAGTTGCA 655
Qy 843 ATTCGTGCA 851
Db 654 ATTCGTGCA 646
RESULT 7
DCU47087/c
LOCUS Daucus carota pathogenesis-related protein mRNA, somatic embryo
DEFINITION
ACCESSION U47087
VERSION U47087.1 GI:1276955
KEYWORDS
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
REFERENCE
AUTHORS Lin,X., Hwang,G.-J. and Zimmerman,J.L.
TITLE Isolation and characterization of a diverse set of genes from carrot somatic embryos
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 727)
AUTHORS Lin,X., Hwang,G.-J. and Zimmerman,J.L.
TITLE Submitted (25-JAN-1996) J. Lynn Zimmerman, Biological Sciences, University of Maryland Baltimore County, 5401 Wilkens Ave, Baltimore, MD 21228, USA
FEATURES Location/Qualifiers

source 1. .727
/organism="Daucus carota"
/mol_type="mRNA"
/strain="Danver Half-long"
/db_xref="taxon:4039"
/clone="Gsa20"
/tissue_type="somatic embryos at the globular stage"
c1. .507
CDS
/codon_start=1
/product="pathogenesis-related protein"
/protein_id="AAB01092.1"
/db_xref="GI:1335877"
/translation="LIFELLISNNILNMGASHLEITSSVSAEKIFXXIVLVDVTV IPKAAPGAYKSDVKGAGGTGATRIITLPEGSPITSMVTVRDANKALTYDSTVIDG DLLEFTESIETHMVVVPADGGSITKTTAIFHTKGDVAVPEENIKFADAQNTALFKAEAYLIAN"
ORIGIN
Query Match 10.6%; Score 90; DB 8; Length 727;
Best Local Similarity 93.7%; Pred. No. 2.6e-07;
Matches 104; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
Qy 742 GGTTTAATAATAGTGTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTT 801
Db 727 GGTTTAATAATAGTGCATATTTTATTTAAAAAGAACNCATACATTACCAGCCAGATGAAGTT 668
Qy 802 ATTCATCACAA - CTCACAAACAAAGTACAAAGAAAAAGTTGCAANTCTGTCA 851
Db 667 ATTCATCACAAAGTTCAACAACAAAGTACAAAGAGAAAGTTGCAANTCTGTCA 617
RESULT 8
E40089/c
LOCUS Plant promoter and terminator.
DEFINITION
ACCESSION E40089
VERSION E40089.1 GI:18627205
KEYWORDS JP 2000166577-A/3.
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 2048)
AUTHORS Nishikawa,S. and Oeda,K.
TITLE Plant promoter and terminator
JOURNAL Patent: JP 2000166577-A 3 20-JUN-2000;
COMMENT SUMITOMO CHEM CO LTD
OS Daucus carota L.
PN JP 2000166577-A/3
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PR SATOMI NISHIKAWA,KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10// (C12N5/10,C12R1:91), PC C12N15/00,
PC C12N5/00, (C12N5/00,C12R1:91)
CC
FT Key Location/Qualifiers
FT promoter (1) . (2048).
FEATURES Location/Qualifiers
source 1. .2048
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 10.4%; Score 88.6; DB 6; Length 2048;
Best Local Similarity 59.6%; Pred. No. 3.9e-07;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;
Qy 345 AATAAGGAGATACGGGGATACGTCAAAGTATTTTTTACGACACAAATAGAAAGCGAGAA 404
Db 342 AATTATTAGATTTGTTGACACGTTAAAAAATAACCCCTGTTTCACTATTGTTATCTCTAAA 283

| | | | |
|----|-----|---|-----|
| Qy | 405 | AGGNTAAAGTGCCTCTATGCAATGGTGAGGAATTGGACGGTTTTAGCTTAGCTTAAAAA | 464 |
| Db | 282 | ATTGAAAGATACATTATTAAGAATGATCTTTTATTGAGCGCGTTTGGGTTAGCTTAAAGAA | 223 |
| Qy | 465 | GTGACTCTTCTACTTTGAAGTAATGAAGTGGAGTAGAACTGATAAAGTAAAGTAATAATTATA | 524 |
| Db | 222 | GTGATTTCTTGCTTATAGCAAAAGAGTGAGGTAGAAATGAGAAGTAA | 173 |
| Qy | 525 | AGTTATTAAAGTGTTTGGAAAGAAATAGAGTGTGAAGAAAGTTAGCATTTTCTACT | 584 |
| Db | 172 | AGTTAATAAAGTGTTTGGAAAGAAAGCGGAAGTTGTGAGAGA | 131 |
| Qy | 585 | TCCAACTTATTTCTCACGACTTCTTAAAAGTACTTCTTACTTTTTTACAAACGGGTCA | 644 |
| Db | 130 | --GAAGTTAGTATTGCGAGCTTTTAAAAATATTTC-TACTTCTTTATACAAACAGGTCA | 74 |
| Qy | 645 | AGGAAAGTGGAGCAAAAGCTGGAGTTACTTCTTATAGAGATGTTTATCTAAATGAGA | 704 |
| Db | 73 | AAAGAAGAAGTGCAGAGACAGCTTCTCTCAACCAACAGGCCCTATGTGCTG | 14 |
| Qy | 705 | AATGACAAACA | 715 |
| Db | 13 | TAGGCACACA | 3 |

| | | | | |
|------------|---|---------|--------|-----------------|
| RESULT 9 | E40090 | 2048 bp | linear | PAT 31-JAN-2002 |
| LOCUS | E40090/c | | | |
| DEFINITION | Plant promoter and terminator. | | | |
| ACCESSION | E40090 | | | |
| VERSION | E40090.1 GI:18627206 | | | |
| KEYWORDS | JP 2000166577-A/4. | | | |
| SOURCE | unidentified | | | |
| ORGANISM | unclassified. | | | |
| REFERENCE | 1 (bases 1 to 2048) | | | |
| AUTHORS | Nishikawa, S. and Oeda, K. | | | |
| TITLE | Plant promoter and terminator | | | |
| JOURNAL | Patent: JP 2000166577-A 4 20-JUN-2000; | | | |
| COMMENT | SUMITOMO CHEM CO LTD | | | |
| | OS Daucus carota L. | | | |
| | PN JP 2000166577-A/4 | | | |
| | PD 20-JUN-2000 | | | |
| | PF 01-OCT-1999 JP 1999281475 | | | |
| | PR | | | |
| | PI SATOMI NISHIKAWA, KENJI OEDA | | | |
| | PC C12N15/09,A01H5/00,C12N1/21,C12N5/10//((C12N5/10,C12R1:91), PC | | | |
| | C12N15/00, | | | |
| | PC C12N5/00, (C12N5/00,C12R1:91) | | | |

| | | | | | |
|----------|-----|---|---------------------------------------|--|--|
| <hr/> | | | | | |
| FEATURES | F1 | promoter | (17.. : (2048)). | | |
| source | | Location/Qualifiers | | | |
| | | 1..2048 | | | |
| | | /organism="unidentified" | | | |
| | | /mol_type="genomic DNA" | | | |
| | | /db_xref="taxon:32644" | | | |
| <hr/> | | | | | |
| ORIGIN | | | | | |
| | | Query Match | 10.4%; Score 88.6; DB 6; Length 2048; | | |
| | | Best Local Similarity | 59.6%; Fred. No. 3.9e-07; | | |
| | | Matches 221; Conservative | 0; Mismatches 119; Indels 31; Gaps 3; | | |
| | | | | | |
| Qy | 345 | AATAAGGAGAAATACGGGATACGTCAAGTATTTTTACGCACACAATAGAAAAGCGAGAA | 404 | | |
| Db | 342 | AATTATTAGAATTTGTTGCACAGCTTA AAAAATAAACCCCTGTTTCACGTATTGTTATCCTAAA | 283 | | |
| Qy | 405 | AGGATAAAGTGCTCTATGANTGCTGAGGAATTCGGACGGTTTAGGTTAGCTTAAAAAAA | 464 | | |
| Db | 282 | ATTGAAAGATACATTATTAATGAATGATCTTTTATTGAGGCCGCTTGGGTTAGCTTAAAAAGAA | 223 | | |
| Qy | 465 | GTGACTTCTTACTTTGAAGTAAATGAATGAGTGAGTAACTGATAAGTAAAGTAAATTAATTATA | 524 | | |

| | | | |
|----|-----|--|-----|
| Db | 222 | GTGATTCTTCCTTATATGACAAAGAGTGGAGTGAAGAATGGAAGTAA-----AAA | 173 |
| Qy | 525 | AGTTATTAAAGTGTGTGGAAAAGAAATAGAAAGTTGTAAGAAAAAGTTAGCAATTTCTTACT | 584 |
| Db | 172 | AGTTAAATAAAGTGTGTGGAAAAGAGCGGAAGTTGTGAGAGA----- | 131 |
| Qy | 585 | TCCAACTTATTTCTCAGACTTCTTAAAGTACTTCTTACTTTTATACACAAACGGGTCA | 644 |
| Db | 130 | --GAAGTTAGTATTCGCAGCTTTTAAAAATATTC-TACTTCTTTATACAAACAGGTCA | 74 |
| Qy | 645 | AGGAAAGTGGAAAGCAAAAGCTGGAGTTACTTCTTATAAGAAATGTTTATACTAAATGAGA | 704 |
| Db | 73 | AAAGAAGAAGAGTCGAACAGCTTCGTCTCTCAACCAACAGGCCCTATGTGCTG | 14 |
| Qy | 705 | AATGACAAACA | 715 |
| Db | 13 | TAGGCACACA | 3 |

| | | | | | |
|------------|--|---------|-----|--------|-----------------|
| RESULT 10 | E40087 | 2052 bp | DNA | linear | PAT 31-JAN-2002 |
| E40087/c | Plant promoter and terminator. | | | | |
| LOCUS | E40087 | | | | |
| DEFINITION | E40087_1 GI:19627203 | | | | |
| ACCESSION | E40087.1 | | | | |
| VERSION | JP 2000166577-A/1. | | | | |
| KEYWORDS | unidentified | | | | |
| SOURCE | unclassified | | | | |
| ORGANISM | unclassified | | | | |
| REFERENCE | 1 (bases 1 to 2052) | | | | |
| AUTHORS | Nishikawa, S. and Oeda K. | | | | |
| TITLE | Plant promoter and terminator | | | | |
| JOURNAL | Patent: JP 2000166577-A 1 20-JUN-2000; | | | | |
| COMMENT | SUMITOMO CHEM CO LTD OS Daucus carota L. PN JP 2000166577-A/1 PD 20-JUN-2000 PF 01-OCT-1999 JP 1999281475 PR PI SATOMI NISHIKAWA, KENJI OEDA PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1:91), PC C12N15/00, CC C12N5/00,(C12N5/00,C12R1:91) Key Location/Qualifiers FH promoter (1)..(2052). FT Location/Qualifiers 1..2052 | | | | |
| FEATURES | source | | | | |

```

ORIGIN
Query Match      10.4%; Score 88.6; DB 6; Length 2052;
Best Local Similarity 59.6%; Pred. No. 3.9e-07;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

Qy 345 AATAAGCAGAAATACGGGATACGTCAAAGTATTTTCAGACACAATAGAAAAGGCAGAA 404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 342 AATTATTTAGAATTTTGTCACCGCTTAANAATAACCCGTGTTCACTGTTATCTCTAAA 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 405 AGGGATAAAGTGCTCTATGAATGGTAGGAATTTGGGACGGTTTTAGTTAGCTTAAAAAAA 464
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 282 ATTGAAAAGATACATTTATGAATGATCTTTTATTTAGGCCCGCTTTGGGTAGCTTAAAAAGAA 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 465 GTGACTTCTTACTTTGAAGTAGTAAGAAGTGGAGTAGAACTGATAAGTAAGTAATTTATA 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 222 GTGATTTCTTGCCTTATAGCAAGAGATGGAGTAGAATAGAGAGTAA-----AAA 173
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 525 AGTTATTTAAAGTGTTTGGAAAAGAAATAGAAAGTTGTAAAGAAAGTTAGCATTTTCTACT 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 AGTTTAAATAAAGTGTTTGGAAAAGAACGCGAAGTTGTGAGAGA----- 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
RESULT 13
166494/c
LOCUS       166494
DEFINITION  Sequence 14 from patent US 5670367.
ACCESSION   166494
VERSION     166494.1
KEYWORDS    GI:2724471
SOURCE      Unknown.
ORGANISM    Unclassified.
REFERENCE   1 (bases 1 to 7218)
AUTHORS     Dörner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE       Recombinant fowlpox virus
JOURNAL     Patent: US 5670367-A 14 23-SEP-1997;
FEATURES    Location/Qualifiers
             source
               1..7218
               /organism="unknown"
               /mol_type="unassigned DNA"

ORIGIN
Query Match          9.1%; Score 77.8; DB 6; Length 7218;
Best Local Similarity 9.6%; Pred. No. 3.4e-05;
Matches 40; Conservative 219; Mismatches 156; Indels 0; Gaps 0;

QY 155 ATTTATTGAGAAATACAACTTTCAGAAAAGCCAACTCTGCGGTTTCAGGTTCTGGAATT 214
DB 1521 ATCTATTGAGTTTCAAAAAGCGCATGTAGGCATCACTGTAATTACCTATCTATGCAAGT 1462

QY 215 AATTAAGCTTTTACATCATCGAAGTATACATATTTCGAAGCGGAGTGAGGAAGAACT 274
DB 1461 AGTTAAAGAGATAGAAGAAATTTGCTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1402

QY 275 TGAGTACGTCGAATGAGAAAATTCAGCAACACAGTACAGAAAGCTCTGGAAGTATGTT 334
DB 1401 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1342

QY 335 GAGCATGTACAAATGAGAGAAATACGGGATACGTCGAAAGTATTTTACGACAAATAGA 394
DB 1341 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1282

QY 395 AAGCGCAGAAAGGATAAGTCTCTATGAATCGTGAGGAATTTGGACGCTTTAGGTAG 454
DB 1281 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1222

QY 455 CTTAAAAAAGTGACTTCTTACTTGAGTAATCAAGTGGAGTACAGTAAAGTAAAGT 514
DB 1221 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1162

QY 515 AATAAATTATAGATTATTAAGTCTTTGGAAGAAATAGAGATTGTAAAGAAAG 569
DB 1161 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1107

DCA18706      8005 bp      DNA      linear      PLN 10-FEB-1999
LOCUS       Dauscus carota Inv*Dc5 gene.
DEFINITION  Y18706
ACCESSION   Y18706
VERSION     Y18706.1
KEYWORDS    beta-fructofuranosidase; Inv*Dc5 gene; invertase; isoform.
SOURCE      Dauscus carota (carrot)
ORGANISM    Dauscus carota
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; campanulids; Apiales; Apiaceae; Apioideae; Scandiceae;
            Daucinae; Dauscus.

REFERENCE   1
AUTHORS     Sturm, A.
TITLE       Molecular characterisation and functional analysis of
            sucrose-cleaving enzymes in carrot (Dauscus carota L.)
JOURNAL     J. Exp. Bot. 47, 1187-1192 (1996)
REFERENCE   2 (bases 1 to 8005)

AUTHORS      Sturm, A.
TITLE        Direct Submission
JOURNAL      Submitted (27-JAN-1999) A. Sturm, Friedrich Miescher-Institut,
Maulbeerstr. 66, CH-4058 Basel, SWITZERLAND
COMMENT      Related sequence X67163.
FEATURES     Location/Qualifiers
             source
               1..8005
               /organism="Dauscus carota"
               /mol_type="genomic DNA"
               /cultivar="Nantaise"
               /db_xref="taxon:4039"
             gene
               1..8005
               /gene="Inv*Dc5"
             5'UTR
               1..2666
               /gene="Inv*Dc5"
             CDS
               join(2767..3160,3724..3732,4656..5515,5605..5766,
5869..6104,6190..6277,6384..6587)
               /gene="Inv*Dc5"
               /EC_number="3.2.1.26"
               /note="soluble acid invertase"
               /codon_start=1
               /product="beta-fructofuranosidase, isoform II"
               /protein_id="CAA7266.1"
               /db_xref="GI:4454115"
               /db_xref="GOA:Q42722"
               /db_xref="EMBL:Q42722"
               /translation="MEHPITISHYTPLPDGEHSPSLTTTNTAEQSSRRBSLTFFVLLFS
SLIACLWGTWLPFNSGNEAVEKSTVPETVEVAPRGVAEGVSKMSFRRLPALNAE
PAPFPWNSVLISWQRSSPHFQPNQNMNDPNGPLPYKGWYHLFYQYNDGAIWGNKI
VNGHVASDGLIHWKHLPMAMVTDHWYDVGWVGTSGATILPDQIIVMLYTGSTNESVOV
QNLAYPADSPDLLIEWKYPGNVLPVPPGIDFKDFRPTTAWRTPEGKRWLIIGSK
LNTGILSVLVDVTKNFTLLDGLVLHVGHTGMECVDFYPSVKFGENGLDTSFDGVG
VHKVWKASLDDDRNDYAIAGTYDVPVSGWVPDNPDELGVIGLRYDYGIVYASKTFYDS
NKKRVLSWIKETDSEISDVRKGNASVQGIPTILPDPKTSNLLQWPEEVNKLRL
NKTFFENVISINTGALPLGEISGQDLDTAEFEVDKESLERVOETNEVDCKRNGGSS
GSGVLPGRGKLTMRILVDHSIVESFSGQGRCTCITSRVYPTKAIYNNAKVFLFPNNAE
ARIASLNIWQNTAQRQTHFADLVI"
               2767..3160
               /gene="Inv*Dc5"
               /number=1
             intron
               3161..3723
               /gene="Inv*Dc5"
               /number=1
             exon
               3724..3732
               /gene="Inv*Dc5"
               /number=2
             intron
               3733..4655
               /gene="Inv*Dc5"
               /number=2
             exon
               4656..5515
               /gene="Inv*Dc5"
               /number=3
             intron
               5516..5604
               /gene="Inv*Dc5"
               /number=3
             exon
               5605..5766
               /gene="Inv*Dc5"
               /number=4
             intron
               5767..5868
               /gene="Inv*Dc5"
               /number=4
             exon
               5869..6104
               /gene="Inv*Dc5"
               /number=5
             intron
               6105..6189
               /gene="Inv*Dc5"
               /number=5
             exon
               6190..6277
               /gene="Inv*Dc5"
               /number=6
             intron
               6278..6383
               /gene="Inv*Dc5"
```

exon /number=6
6384..6587
/gene="Inv*Dcs"
3' UTR /number=7
6588..8005
/gene="Inv*Dcs"
ORIGIN

| | | | | |
|-----------------------|-----------------|---|------------|--------------|
| Query Match | 9.0%; | Score 76.4; | DB 8; | Length 8005; |
| Best Local Similarity | 66.9%; | Pred. NO. 6.2e-05; | | |
| Matches 170; | Conservative 0; | Mismatches 51; | Indels 33; | Gaps 3; |
| Qy | 425 | ATGGTGAAGATTGGACGGTTAGGTTAGCTTAAAGAAAGTGACTTCTTCTTCAAGTA | 484 | |
| Db | 1040 | AGGGTCAGTCATAGGGGCGTTTGGGTTAGCTTAAAGAAAGTGATTTCTTCTTAAATTA | 1099 | |
| Qy | 485 | ATGAAGTGGAGTAGAAGCTGATAAGTAAAGTAATAATTAAGTTATTAAAGTGTGGAA | 544 | |
| Db | 1100 | GAAGAAGTAGAGTAGAAGTGAGAGTAA-----ATAAATTAATAAGTGTGGAA | 1149 | |
| Qy | 545 | AAGAAATAGAGTTGT- AAAGAAAGTTAGCATTTTCTTCTTCTTCAAGTATTTCTCAGA | 603 | |
| Db | 1150 | AAGAAAGCAAGCTGTGAGAGAGAGTAGCATTTCTCAACTT----- | 1191 | |
| Qy | 604 | CTTCTTAAAGTACTTCTTACTTTTACACAAAGCGGTCAAGGAAAGTGGACCAAAA | 663 | |
| Db | 1192 | ----TTTAAATAGTCTTCTGCTTATTACACAAAGCGGTCAAGGAAAGCAGACGGA | 1247 | |
| Qy | 664 | GCTGGAGTTACTTC 677 | | |
| Db | 1248 | GCAGCTCTACTTC 1261 | | |

RESULT 15
DCU56392/c
LOCUS DCU56392 2215 bp DNA linear PLN 01-FEB-1999
DEFINITION Daucus carota elicitor induced glycoprotein iEP4 gene, complete cds.
ACCESSION U56392
VERSION U56392.1 GI:4204869
KEYWORDS
SOURCE Daucus carota (carrot)
ORGANISM Daucus carota
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; campanulids; Apiales; Apiaceae; Apioidae; Scandiceae; Daucinae; Daucus.
REFERENCE 1 (bases 1 to 2215)
Bertinetti, C.M. and Ugalde, R.A.
Studies on the response of carrot cells to a Sclerotinia sclerotiorum elicitor. Induction of expression of an extracellular glycoprotein mRNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2215)
Bertinetti, C.M. and Ugalde, R.A.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (23-APR-1996) C.M. Bertinetti, Instituto de Investigaciones Bioquimicas 'Fundacion Campomar', Patricias Argentinas 435, Capital Federal, Buenos Aires 1405, Argentina
FEATURES
source
1..2215
/organism="Daucus carota"
/mol_type="genomic DNA"
/db_xref="taxon:4039"
/cell_lines="ZA+ line"
join(55..292,365..580,1077..1246,1358..1628,1797..2077)
join(55..292,365..580,1077..1246,1358..1628,1797..2077)
/note="glycoprotein"
/codon_start=1
/product="iEP4"
/protein_id="AAD11468.1"
/db_xref="GI:4204870"
/translation="MGKHSAPSVFLFSLVAHSHVSASSQTCDFPPIAFNFGDANSDTG

ORIGIN
Query Match 8.8%; Score 74.8; DB 8; Length 2215;
Best Local Similarity 70.4%; Pred. No. 0.00016;
Matches 114; Conservative 0; Mismatches 47; Indels 1; Gaps 1;
Qy 437 TGGACCGTTTAGGTTAGCTTAAAGAAAGTGACTTCTTCTTGAAGTAATGAAGTGGAGT 496
Db 871 TGGGCGCGTTTGGGTAACCTTTAAAAAAGT-ACCTTCTTCTTAAAGTAAAGAGTGAAC 813
Qy 497 AGAACTGATAAGTAAAGTAATAATTATAAGTTATTAAAGTGTGGAAAGAAATAGAA 556
Db 812 AGAAGTAAAGTGTAAATATTAGAACTTATAAGCGATTAAACTGTTTGGGAAACAGCAGAAAT 753
Qy 557 TTGTAAGAAAGTGAAGTGAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 598
Db 752 CCTGAAACAAAGTGAAGTGAAGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 711
Search completed: December 6, 2004, 22:47:31
Job time : 3918.95 secs

APAAWFGNPPFPQSYFNGSAGRVSDGRLLIDFMATDGLPFLHPYMDSLGANFSGH
ANFANILLSTIALPTSNIIIGVRPFRGLNPVNLIDIQVAFQFVNRSTQGEAFNFMF
KQDYFSQALYTLIDIGQIDITQBFITNKTDDIIRKAVFGLISSLSNIIQIYSLGSRF
WIHNLGPNGLPILLTLAPVDPDQLDSAGCAKRYNDLTQYFNSLKKGVQDLQDLPL
AATYVDVYAKYSLYQEPKAYGFTHPLETCCGFGGRYNGEPLSCGSTITVNGTQLT
VGPCENPAEYINYEQTYYQAADQITFNKISTGELSPPNSLKTACPKLSLPRVSDI"

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 15:16:44 ; Search time 470.618 Seconds
(without alignments)
9492.325 Million cell updates/sec

Title: US-09-806-197-2

Perfect score: 851

Sequence: 1 ctgaaaggaagttcatcg.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04.*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002as.*
- 7: geneseqn2002bs.*
- 8: geneseqn2003as.*
- 9: geneseqn2003bs.*
- 10: geneseqn2003cs.*
- 11: geneseqn2003ds.*
- 12: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 851 | 100.0 | 851 | 3 | Aaa37960 Carrot te |
| 2 | 123.6 | 14.5 | 739 | 2 | Aag92559 cDNA clon |
| 3 | 123.6 | 14.5 | 739 | 3 | Aaa37965 Plasmid # |
| 4 | 88.6 | 10.4 | 2048 | 3 | Aaa37962 Plasmid # |
| 5 | 88.6 | 10.4 | 2048 | 3 | Aaa37963 Plasmid # |
| 6 | 88.6 | 10.4 | 2052 | 3 | Aaa37959 Carrot pr |
| 7 | 88.6 | 10.4 | 2052 | 3 | Aaa37961 Carrot pr |
| 8 | 88.6 | 10.4 | 2056 | 3 | Aaa37964 Plasmid # |
| 9 | 71 | 8.3 | 2042 | 2 | Aav15144 New promo |
| 10 | 63.6 | 7.5 | 836 | 2 | Aav15146 plant gen |
| 11 | 62 | 7.3 | 2831 | 10 | Adc56759 Carrot DN |
| 12 | 62 | 7.3 | 2865 | 10 | Adc56761 Carrot DN |
| 13 | 54.6 | 6.4 | 2114 | 6 | Abk40080 Human che |
| 14 | 52.8 | 6.2 | 4045 | 6 | Ab192321 Chemical |
| 15 | 52.8 | 6.2 | 4045 | 6 | Aad22342 Chemical |
| 16 | 52.2 | 6.1 | 8056 | 8 | Abz10246 Haematopo |
| 17 | 52 | 6.1 | 539 | 5 | Ad138243 Human ova |
| 18 | 52 | 6.1 | 539 | 5 | Ad173110 Human ova |
| 19 | 52 | 6.1 | 13712 | 6 | Ab133531 Human imm |
| 20 | 51.8 | 6.1 | 19787 | 6 | Ab133451 Human imm |
| 21 | 51.6 | 6.1 | 8676 | 6 | Ab170452 Chemical |

| | | | | | | |
|------|------|-----|--------|----|----------|-----------|
| C 22 | 51.6 | 6.1 | 8676 | 6 | AAS61415 | Human gen |
| C 23 | 49.8 | 5.9 | 9157 | 6 | ABL33500 | Human imm |
| C 24 | 49.6 | 5.8 | 17721 | 6 | ABL33728 | Human imm |
| C 25 | 49.2 | 5.8 | 113515 | 6 | ABL34175 | Human imm |
| C 26 | 49 | 5.8 | 9789 | 2 | AAT41852 | cDNA enco |
| C 27 | 48.4 | 5.7 | 8404 | 4 | AAS46500 | Tumour su |
| C 28 | 48.4 | 5.7 | 8404 | 6 | ABL33595 | Human imm |
| C 29 | 48 | 5.6 | 6378 | 6 | ABL32176 | Human imm |
| C 30 | 48 | 5.6 | 6378 | 6 | ABO67027 | Human ang |
| C 31 | 47.8 | 5.6 | 14924 | 6 | ABL32225 | Human imm |
| C 32 | 47.8 | 5.6 | 14924 | 6 | ABL54322 | Chemical |
| C 33 | 47.8 | 5.6 | 19380 | 6 | AAS61427 | Human gen |
| C 34 | 47.8 | 5.6 | 29993 | 10 | ADB37663 | Human che |
| C 35 | 47.4 | 5.6 | 34548 | 6 | ABL70604 | Chemical |
| C 36 | 47.4 | 5.6 | 615 | 2 | ADR01494 | A. Gossyp |
| C 37 | 47.4 | 5.6 | 7195 | 4 | AAS45325 | Chemical |
| C 38 | 47.4 | 5.6 | 7195 | 6 | ABK28166 | DNA trans |
| C 39 | 47.4 | 5.6 | 9180 | 6 | ABL33965 | Human imm |
| C 40 | 47.2 | 5.5 | 875 | 3 | AAA01920 | Human col |
| C 41 | 46.8 | 5.5 | 5689 | 4 | AAS45384 | Chemical |
| C 42 | 46.8 | 5.5 | 5689 | 4 | AAS46426 | Tumour su |
| C 43 | 46.8 | 5.5 | 5689 | 6 | ABK28226 | DNA trans |
| C 44 | 46.6 | 5.5 | 40862 | 6 | ABL34073 | Human imm |
| C 45 | 46.4 | 5.5 | 5145 | 6 | ABL32349 | Human imm |

ALIGNMENTS

RESULT 1

AAA37960

ID AAA37960 standard; DNA; 851 BP.

AC AAA37960;

DT 18-AUG-2000 (first entry)

DE Carrot terminator sequence #2.

KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.

OS Daucus carota.

PN WO200020613-A1.

PD 13-APR-2000.

PF 28-SEP-1999; 99WO-JP005303.

PR 02-OCT-1998; 98JP-00281124.

PA (SUMO) SUMITOMO CHEM CO LTD.

PI Nishikawa S, Oeda K;

DR WPI; 2000-303791/26.

PT New plant promoters and terminators from Daucus carota L., useful in plant breeding, for e.g. controlling fertilities of plants.

PS Claim 2; Page 69-71; 81pp; English.

CC This sequence represents a carrot terminator sequence. The invention relates to plant promoters and terminators from Daucus carota L. which are capable of expressing a gene of interest in plants. The invention also includes a chimeric gene characterized in that it comprises the promoter and a desired gene linked to each other in the form capable of functioning. A method of producing a transformant comprising introducing the promoter, the chimeric gene or a vector comprising the promoter and a desired gene or terminator sequence into a host cell. The plant promoters and terminators are useful in plant breeding, for e.g. fertilities of plants may be controlled by expressing, in the host cells, a sense or antisense gene of a male sterility related gene such as S-locus-specific


```

AAA37965/c
ID AAA37965 standard; DNA; 739 BP.
XX
AC AAA37965;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #4 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
XX plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 76-78; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
XX introducing a mutation into a carrot promoter. The invention relates to
XX plant promoters and terminators from Daucus carota L. which are capable
XX of expressing a gene of interest in plants. The invention also includes a
XX chimeric gene characterized in that it comprises the promoter and a
XX desired gene linked to each other in the form capable of functioning. A
XX method of producing a transformant comprises introducing the promoter,
XX the chimeric gene or a vector comprising the promoter and a desired gene
XX or terminator sequence into a host cell. The plant promoters and
XX terminators are useful in plant breeding, for e.g. fertilities of plants
XX may be controlled by expressing, in the host cells, a sense or antisense
XX gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 739 BP; 218 A; 157 C; 151 G; 213 T; 0 U; 0 Other;

Query Match 14.5%; Score 123.6; DB 3; Length 739;
Best Local Similarity 96.9%; Pred. No. 1.2e-17;
Matches 126; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 722 TGAGATGAATATGATTTGTTTAAATAGTGTATTTTATTTAAAGATCGCATAC 781
Db 718 TGAGATGAATATGATTTGTTTAAATAGTGTATTTTAAAGATCGCATAC 659

QY 782 ATTACCAGCAGATGAATTTATTCACCACTCACAACAAAGTACAAAGAAAGTTGC 841
Db 658 ATTACCAGCAGATGAATTTATTCATCAAGTACCAACAAAGTACAAAGAAAGTTGC 599

QY 842 AATTCGTGCA 851
Db 598 AATTCGTGCA 589

RESULT 4
ID AAA37962/c
XX AAA37962 standard; DNA; 2048 BP.
XX
AC AAA37962;
XX
DT 18-AUG-2000 (first entry)
XX

```

```

DE Plasmid #1 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
XX plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 71-73; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
XX introducing a mutation into a carrot promoter. The invention relates to
XX plant promoters and terminators from Daucus carota L. which are capable
XX of expressing a gene of interest in plants. The invention also includes a
XX chimeric gene characterized in that it comprises the promoter and a
XX desired gene linked to each other in the form capable of functioning. A
XX method of producing a transformant comprises introducing the promoter,
XX the chimeric gene or a vector comprising the promoter and a desired gene
XX or terminator sequence into a host cell. The plant promoters and
XX terminators are useful in plant breeding, for e.g. fertilities of plants
XX may be controlled by expressing, in the host cells, a sense or antisense
XX gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2048;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

QY 345 RATAAGGAGATACGGGATACGTCAAGTATTTTACGACACATAGAAAGGCAGAA 404
Db 342 AATTATTAGATTTGTGACACGTTAAATAAACCCCTGTTCCGTATTGTTATCTCTAAA 283

QY 405 AGGGATAAAGTCTCTATGAATGGTGAGGAATTTGGACCGTTTAGTTAGCTTAAAAAAA 464
Db 282 ATTGAAAGATACATTTATGATGATCTTTTATTGAGCGCTTTGGTTAGCTTAAAGAA 223

QY 465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAAGTAAAGTAAAGTAAATATA 524
Db 222 GTGATTTCTTGTCTTATAGCAAGAAGTGGAGTAGAAGTAAAGTAAAGTAAAGTAA 173

QY 525 AGTTTAAAGTGTGGTGGAAAGAAATAGAGTCTTAAGAAAGTAAAGTAAAGTAAAGT 584
Db 172 AGTTTAAAGTGTGGTGGAAAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 131

QY 585 TCCAACTTATTCTCACCAGCTTCTTAAAGTACTTCTTACTTTTAAACAAACGGGTCA 644
Db 130 --GAAGTTAGTATTCGACGCTTTTAAATAATATTC-TACTTCTTATACAAACAGGTCA 74

QY 645 AGGAAAGTGAAGCAAAAAGCTGGAGTTACTTCTTATAGAAGTGTATTAATAATAGA 704
Db 73 AAAGAAGAAAGAGTCAAGAGCAGCTTCTCTCTCAACCAAAACAGGCCCTATGTGCTG 14

QY 705 AATGACAACA 715
Db 13 TAGGGCACACA 3

```

```

RESULT 5
AAA37963/c
ID AAA37963 standard; DNA; 2048 BP.
XX
AC AAA37963;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #2 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 73-74; 8lpp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2048;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

QY 345 AATAAGGAGATACGGGGATACGTCACAAAGTATTTTACGACACAAATAGAAAAGCAGAA 404
DB 342 AATTATTAGAAATTTGTCACAGTTTAAATAAATACCTGTTTACGATTTATCTTAA 283
QY 405 AGGGATAAAGTCTCTATGAATGGTGAAGAAATGGGACGGTTTAGTTAGCTTAAAAAAA 464
DB 282 ATTGAAGAATACATTTATGAATGATCTTTTATTGAGGCGGTTTGGGTTAGCTTAAAGAA 223
QY 465 GTGACTTCTTACTTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 524
DB 222 GTGATTTCTTCTTATAGCAAGAAAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 173
QY 525 AGTTATTAAAGTGTGTTGAAAAAGAAATAGAAAGTGTGTTTAAAGAAAGTGTAGCATTTTCTACT 584
DB 172 AGTTAATAAGTGTGTTGAAAAAGAAAGTGTGTTTAAAGAAAGTGTGTTTAAAGTGTGTTTAAAGT 131
QY 585 TCCAACTATTCTCAGCACTCTTAAAGTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 644
DB 130 --GAAGTTAGTATTCGCAGCTTTTAAAAAATATTTTC-TACTTCTTTTATACAAACAGGTCA 74
QY 645 AGGAAAGTGAAGCAAAAAGCTGGAGTTACTTCTTTAAGAATGTTTATACTACTAATAGAGA 704

```

```

DB 73 AAAGAAGAAGTCAAGCAGCTTCTCTCAACCAACAGCCCTATGTGTG 14
QY 705 AATGACAAACA 715
DB 13 TAGGGCACACA 3

RESULT 6
AAA37959/c
ID AAA37959 standard; DNA; 2052 BP.
XX
AC AAA37959;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #1.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 69-70; 8lpp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2052;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

QY 345 AATAAGGAGATACGGGGATACGTCACAAAGTATTTTACGACACAAATAGAAAAGCAGAA 404
DB 342 AATTATTAGAAATTTGTCACAGTTTAAATAAATACCTGTTTACGATTTATCTTAA 283
QY 405 AGGGATAAAGTCTCTATGAATGGTGAAGAAATGGGACGGTTTAGTTAGCTTAAAAAAA 464
DB 282 ATTGAAGAATACATTTATGAATGATCTTTTATTGAGGCGGTTTGGGTTAGCTTAAAGAA 223
QY 465 GTGACTTCTTACTTGAAGTAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 524
DB 222 GTGATTTCTTCTTATAGCAAGAAAGTGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 173
QY 525 AGTTATTAAAGTGTGTTGAAAAAGAAATAGAAAGTGTGTTTAAAGAAAGTGTAGCATTTTCTACT 584

```

```

Db      172 AGTTAATAAAGTGTGGGAAAGAGCGGAAGTTGTGAGAGA----- 131
QY      585 TCCAACTATTCTCAGCACTCTTAAAGTACTTCTTACTTTTACACAAACGGGTCA 644
Db      130 --GAAGTTAGTATTCGAGCTTTTAAATATATTC-TACTTCTTTATACAAACAGGTCA 74
QY      645 AGGAAAGTGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAAATGTTTATATACTAAATGAGA 704
Db      73 AAAGAAGAAGAGTCAAGAGCAGCTTCTGCTCTCTCAACCAACAGGCGCTATGTGCTG 14
QY      705 AATGACAAACA 715
Db      13 TAGGGCACACA 3

RESULT 7
AAA37961/c
ID AAA37961 standard; DNA; 2052 BP.
XX
AC AAA37961;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #2.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 78-79; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other;

Query Match 10.4%; Score 88.6; DB 3; Length 2052;
Best Local Similarity 59.8%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

QY      345 AATAAGGAGAAATACGGGGATACGTCAAAGTATTTTACGACACATAAGAAAGGCAGAA 404
Db      342 AATTATTAGAAATTTGTTGACACGCTTAAATAACCTGTTCACTGTTATGTTATCTAAA 283
QY      405 AGGGATAAAGTGCTCTATGAATGGTGGAGGAATGGGACGGTTTGTAGCTTAAAGAAA 464
Db      282 ATTGAAGAATACATTATGAATGATCTTTTATTGAGGCGCTTTGGGTTAGCTTAAAGAA 223

```

```

QY      465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
Db      222 GTGATTTCTTGTCTTAGCAAGAAGTGGAGTAGAAATGAGAAGTAA-----AAA 173
QY      525 AGTTATTAAGTGTGTTGGAAAAAGAAATAGAAGTTGTFAAAGAAAAGTTAGCATTTTCTACT 584
Db      172 AGTTAATAAAGTGTGTTGGAAAAAGAGCGGAAGTTGTGAGAGA----- 131
QY      585 TCCAACTATTCTCAGCACTCTTAAAGTACTTCTTACTTTTACACAAACGGGTCA 644
Db      130 --GAAGTTAGTATTCGAGCTTTTAAATAATATTC-TACTTCTTTATACAAACAGGTCA 74
QY      645 AGGAAAGTGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAAATGTTTATATACTAAATGAGA 704
Db      73 AAAGAAGAAGAGTCAAGAGCAGCTTCTGCTCTCAACCAACAGGCGCTATGTGCTG 14
QY      705 AATGACAAACA 715
Db      13 TAGGGCACACA 3

```

```

RESULT 8
AAA37964/c
ID AAA37964 standard; DNA; 2056 BP.
XX
AC AAA37964;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;

```

```

Query Match 10.4%; Score 88.6; DB 3; Length 2056;
Best Local Similarity 59.6%; Pred. No. 7.3e-10;
Matches 221; Conservative 0; Mismatches 119; Indels 31; Gaps 3;

```

QY 345 AATAGGAGATACGGGATAGCTCAAGATATTTTTTACGACACATAGAAAAAGCAGAA 404
DB 342 AATTATTAGAATTTGTTGACACGTTTAAAAAATAACCCCTGTTTACGATATTGTTATCCTAAA 283
QY 405 AGGGATAAAGTGCTCTATGAATGGTGAGGAATTTGGGACGGTTTAGCTTTAGCTTTAAAAAAA 464
DB 282 ATTGAAGATACATTTATGAATGATCTTTTATGAGCGCGTTTGGGTTAGCTTTAAAGAA 223
QY 465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAACTGATAAGTAAAGTAATAATTATA 524
DB 222 GTGATTTCTTGCTTATAGCAAGAGTGGAGTAGAATAGAGTAA-----AAA 173
QY 525 AGTTATTAAGTGTGTTGAAAAAGAAATAGAGTGTGTAAGAAAGTTAGCATTTTCTACT 584
DB 172 AGTTAATAAGTGTGTTGAAAAAGAGCGGAAGTTGTGAGAGA----- 131
QY 585 TCCAACTTATTCTCAGCACTTCTTAAAGTACTTCTTACTTTTACACAAACGGGTCA 644
DB 130 --GAAGTTAGTATTGCGAGCTTTTAAAAAATATTTC-TACTTCTTTATACAAACAGGTCA 74
QY 645 AGGAAAGTGGAGCAAAAAAGCTGGAGTTACTTCTTATAAGAAATGTTTATCTAATAATGAGA 704
DB 73 AAAGAAGAAGAGTCAAGAGCAGCTTCTGCTTCTCAACCAACAGCGCCCTATGTGCTG 14
QY 705 AATGACAACA 715
DB 13 TAGGGCACACA 3

RESULT 9

AAV15144/c

ID AAV15144 standard; DNA; 2042 BP.

AC AAV15144;

DT 02-JUL-1998 (first entry)

DE New promoter used for root-specific expression in plants.

XX Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW 86.

OS Daucus carota.

PN EP824150-A2.

PD 18-FEB-1998.

XX 12-AUG-1997; 97EP-00113923.

PR 12-AUG-1996; 96JP-00212680.

PA (SUMO) SUMITOMO CHEM CO LTD.

PI Torikai S, Oeda K;

DR WPI; 1998-122310/12.

XX New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.

PS Claim 2; Page 15-16; 31pp; English.

CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants

SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;
Query Match 8.3%; Score 71; DB 2; Length 2042;
Best Local Similarity 73.6%; Pred. No. 5.5e-06;
Matches 117; Conservative 0; Mismatches 40; Indels 2; Gaps 2;

QY 428 GTGAGGAATTGGGACGGTTTAGCTTAAAGTAAAGTCTTCTTACTTGAAGTAATG 487
DB 1159 GGGATGAGAAAGGTCGTTTGGTTAAACTTAAAAAAGTGTC-TCTTGCAATAAGTAAAG 1101
QY 488 AAGTGGAGTAGAACTGATAAGTAAAGTAATAATTAAGTTATTAAAGTGTTTGGAAAAAG 547
DB 1100 AAGTGCACCAAGTAGAAGTAATTAAGATTATTAAGTGAATTAATTTGTTGGGAAG 1041
QY 548 AATAGAAAGTTGT-AAAGAAAAAGTTAGCATTTTCTACTT 585
DB 1040 AAGCAGAAGTCGTGAAACTAAAGTTAGTATTGTTCGGTTT 1002

RESULT 10

AAV15146

ID AAV15146 standard; DNA; 836 BP.

AC AAV15146;

DT 02-JUL-1998 (first entry)

DE Plant gene terminator sequence.

XX Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
KW terminator; ds.

OS Daucus carota.

PN EP824150-A2.

PD 18-FEB-1998.

XX 12-AUG-1997; 97EP-00113923.

PR 12-AUG-1996; 96JP-00212680.

XX (SUMO) SUMITOMO CHEM CO LTD.

PA Torikai S, Oeda K;

DR WPI; 1998-122310/12.

XX New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.

PS Claim 25; Page 20; 31pp; English.

CC The present sequence represents a terminator isolated from a carrot
CC genomic library of cultivar Kuroda Gosun. The specification describes a
CC novel promoter, also isolated from the same library. The novel promoter
CC can be used to direct root-specific expression in plant cells. Since the
CC promoter enables expression of a desired protein in the roots of a plant,
CC it is useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants

SQ Sequence 836 BP; 278 A; 122 C; 141 G; 295 T; 0 U; 0 Other;

Query Match 7.5%; Score 63.6; DB 2; Length 836;

Best Local Similarity 71.2%; Pred. No. 0.0002;

Matches 84; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 437 TGGGACGGTTTAGCTTAAAGTAAAGTCTTCTTACTTGAAGTAATGAAGTGGAGT 496

DB 518 TGGGGCAGTTTGGCTGGACTTAAAAAAGTGAATTTGCTTAAATAATAAGTAGATT 577

XX 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX (EPiG-) EPIGENOMICS AG.
PI Olek A, Piepenbrock C, Berlin K;
XX WPI; 2002-154757/20.
DR
XX New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated with
PT pharmacogenomics and for therapy of diseases e.g. cancer.
XX Claim 1; SEQ ID NO 162; 24pp; English.
XX The invention relates to a nucleic acid comprising a sequence at least 18
CC bases in length of a segment of the chemically pretreated DNA of genes
CC associated with pharmacogenomics according to one of the sequences of the
CC genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1 (NM_000497), CYP3A3
CC (NM_000776 and NM_017460), DPXD (NM_000110), EPHX2 (NM_000497), OCLN
CC (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360), MRF (NM_004996,
CC NM_019900, NM_019901, NM_019902, NM_019862, NM_019898, NM_019899) and
CC their complementary sequences, or a sequence (S1) chosen from 87
CC sequences and their complements. The chemical pretreatment is bisulphite
CC treatment to convert cytosines (but not methyl-cytosines) into uracils.
CC Also included are an oligomer (II) in particular an oligonucleotide or a
CC peptide nucleic acid (PNA)-oligomer, comprising in each case at least one
CC base sequence having a length of 9 nucleotides which hybridises to or is
CC identical to a chemically pretreated DNA of genes associated with
CC pharmacogenomics and their complements, arranged in an array for
CC analysing diseases associated with the methylation state (CPG) and/or
CC detecting SNPs (single nucleotide polymorphisms) of the 87 sequences. The
CC oligomers may also be used as PCR primers. The set of 87 nucleic acids
CC and their complements is useful for diagnosis and therapy of solid
CC tumours and cancer. The present sequence represents one the 87 DNA
CC sequences or its complement. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2114 BP; 504 A; 37 C; 451 G; 1122 T; 0 U; 0 Other;
Query Match 6.4%; Score 54.6; DB 6; Length 2114;
Best Local Similarity 56.4%; Pred. No. 0.023;
Matches 102; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
QY 649 AAGTGGAGCAAAAGCTGGAGTTACTTCTTATAGAAATGTTTACTAAATGAGAAATG 708
DB 1499 AAATAAACAACCAAAAATATATTTTAATCAATTAATTAATTAATTAATTAATTAATTA 1440
QY 709 ACAACACAGAAATGAGAAATGATGATTTGTTTAAATAATAGTGATTTATTTTAA 768
DB 1439 ACAATAATACAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1380
QY 769 AAAGATCCGATACATTACAGCCAGATGAAGTTATTCATCACTCAACAAGATCA 828
DB 1379 AATAAAATCTTACTTTACTACCCAAATTAATAAATAAATAAATAAATAAATAAATAA 1320
QY 829 A 829
DB 1319 A 1319
RESULT 14
ABL92321
ID ABL92321 standard; DNA; 4045 BP.
XX
AC ABL92321;
XX
DT 01-JUL-2002 (first entry)
XX
DB Chemically treated DNA repair gene fragment complementary to#65.

XX DNA repair; cytosine methylation; PMS2L1; PMS2L12; PMS2L3; PMS2;
KW L4; PMS2L5; PMS2L6; MGMT; MSH2; NUDT1; TDG; INPPL1; RFC4; DDT1L; FANCB;
KW XRC8; ataxia telangiectasia; aging; Bloom's syndrome; Cockayne syndrome;
KW Nijmegen breakage syndrome; Werner syndrome; immunodeficiency;
KW trichiodystrophy; Fanconi's anaemia; solid tumour; cancer; ds.
XX Unidentified.
OS
XX WO200181622-A2.
PN
XX 01-NOV-2001.
PD
XX 06-APR-2001; 2001WO-EP003972.
PF
XX 06-APR-2000; 2000DE-01019058.
PR 07-APR-2000; 2000DE-01019173.
PR 30-JUN-2000; 2000DE-01032529.
PR 01-SEP-2000; 2000DE-01043826.
XX (EPiG-) EPIGENOMICS AG.
PA
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2002-034446/04.
DR
XX New nucleic acid derived from genes associated with DNA repair, useful
PT for diagnosis, e.g. of ataxia telangiectasia, by determination of
PT cytosine methylation.
PT
XX Claim 1; SEQ ID NO 130; 25pp + Sequence Listing; English.
PS The invention relates to nucleic acids containing a sequence of at least
XX 18 nucleotides of chemically treated DNA of genes associated with DNA
CC repair, and their complements. The invention also relates to nucleic
CC acids comprising at least 18 base pairs of the chemically pretreated DNA
CC of genes associated with DNA repair selected from PMS2L1, PMS2L12,
CC PMS2L2, PMS2L3, PMS2, L4, PMS2L5, PMS2L6, MGMT, MSH2, NUDT1, TDG, INPPL1,
CC RFC4, DDT1L, FANCB, or XRC8. Nucleic acids of the invention and related
CC oligomers, are useful for diagnosis of diseases associated with gene
CC repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
CC Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome,
CC immunodeficiency, trichiodystrophy, Fanconi's anaemia, solid tumours
CC and cancer, particularly by determining status of cytosine methylation
CC and/or by detecting single-nucleotide polymorphisms. Determination of
CC individual methylation patterns may allow development of individualised
CC therapies. The sequences given in records ABL92192-ABL92335 represent
CC chemically pre-treated DNA fragments from genes associated with DNA
CC repair, and their complements. Note: The sequence data for this patent is
CC not represented in the specification, but is based on sequence
CC information supplied by the European Patent Office
XX
SQ Sequence 4045 BP; 1310 A; 64 C; 694 G; 1977 T; 0 U; 0 Other;
Query Match 5.2%; Score 52.8; DB 6; Length 4045;
Best Local Similarity 47.5%; Pred. No. 0.064;
Matches 188; Conservative 0; Mismatches 207; Indels 1; Gaps 1;
QY 374 TATTTTTTACGACACAATAGAAAAGCGAAGGATAAGTGCTCTATGATGATGAGG 433
DB 858 TTGTGTTTAAAGAAATATATATTTTGTAAAAAAGAAAAAAGAAAGAAAAA 917
QY 434 AATTGGGACGGTTAGGTTAGCTTTAAAAAAGTGAAGTCTTTCTTCTTGAAGTGAAGTGG 493
DB 918 AAGGGGGGAGTGTAGTGGAAATTCAGTATATAAATGTTTAAAGTGTGTTTAAATAATAA 977
QY 494 AGTAGAACTGATAAGTAAAGTAAATATATATAGTTTATTAAGTCTTTGGAAGAAAGTAG 553
DB 978 AATAAACGTAAATGTTTGTAGGGGAGAGTGTGTTTAAAGTATTTGTTAAAGTTAGG 1037
QY 554 AAGTTGTAAGAAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCAGCACTTCTTAAAA 613
DB 1038 AAGAAATAATGGGTATTTTGTGTTAGGTAGGTAGGTAATATTTTAAAAATTTTATAAGTG 1097

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:08:52 ; Search time 85.3031 Seconds
(without alignments)
7090.965 Million cell updates/sec

Title: US-09-806-197-2
Perfect score: 851
Sequence: 1 ctgaaaggaagtctcatcg.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues
Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|---------------------|--------------------|
| C 1 | 123.6 | 14.5 | 739 | 1 US-08-363-010-3 | Sequence 3, Appl |
| C 2 | 77.8 | 9.1 | 7218 | 1 US-08-232-463-14 | Sequence 14, Appl |
| C 3 | 71 | 8.3 | 2042 | 2 US-08-911-434A-2 | Sequence 2, Appl |
| C 4 | 63.6 | 7.5 | 836 | 2 US-08-911-434A-5 | Sequence 5, Appl |
| C 5 | 55.8 | 6.6 | 1141 | 4 US-09-806-708B-22 | Sequence 22, Appl |
| C 6 | 47.4 | 5.6 | 615 | 3 US-08-998-416-186 | Sequence 186, Appl |
| C 7 | 45.2 | 5.3 | 1141 | 4 US-09-806-708B-22 | Sequence 22, Appl |
| C 8 | 44.8 | 5.3 | 1341 | 3 US-08-960-780-51 | Sequence 51, Appl |
| C 9 | 44.8 | 5.3 | 1341 | 3 US-09-073-898-51 | Sequence 51, Appl |
| C 10 | 44.8 | 5.3 | 1341 | 3 US-09-371-913A-2 | Sequence 2, Appl |
| C 11 | 44.8 | 5.3 | 1341 | 4 US-09-307-106-11 | Sequence 11, Appl |
| C 12 | 44.8 | 5.3 | 1341 | 4 US-09-967-805-2 | Sequence 2, Appl |
| C 13 | 44.8 | 5.3 | 1341 | 4 US-09-850-351A-51 | Sequence 51, Appl |
| C 14 | 44.8 | 5.3 | 4041 | 1 US-08-471-033-22 | Sequence 22, Appl |
| C 15 | 44.8 | 5.3 | 4041 | 2 US-08-471-044-22 | Sequence 22, Appl |
| C 16 | 44.8 | 5.3 | 4041 | 2 US-08-463-483A-22 | Sequence 22, Appl |
| C 17 | 44.8 | 5.3 | 4041 | 2 US-08-471-046A-22 | Sequence 22, Appl |
| C 18 | 44.8 | 5.3 | 4041 | 2 US-08-470-566B-22 | Sequence 22, Appl |
| C 19 | 44.8 | 5.3 | 4041 | 2 US-08-469-334-22 | Sequence 22, Appl |
| C 20 | 44.8 | 5.3 | 4041 | 3 US-09-300-529-22 | Sequence 22, Appl |
| C 21 | 44.8 | 5.3 | 6049 | 1 US-08-471-033-1 | Sequence 1, Appl |
| C 22 | 44.8 | 5.3 | 6049 | 2 US-08-471-044-1 | Sequence 1, Appl |
| C 23 | 44.8 | 5.3 | 6049 | 2 US-08-463-483A-1 | Sequence 1, Appl |
| C 24 | 44.8 | 5.3 | 6049 | 2 US-08-471-046A-1 | Sequence 1, Appl |
| C 25 | 44.8 | 5.3 | 6049 | 2 US-08-470-566B-1 | Sequence 1, Appl |
| C 26 | 44.8 | 5.3 | 6049 | 2 US-08-469-334-1 | Sequence 1, Appl |
| C 27 | 44.8 | 5.3 | 6049 | 3 US-09-300-529-1 | Sequence 1, Appl |

| | | | | | |
|------|------|-----|--------|-----------------------|-------------------|
| C 28 | 44.6 | 5.2 | 6866 | 4 US-10-204-708-20 | Sequence 20, Appl |
| C 29 | 43.6 | 5.1 | 1296 | 4 US-09-248-796A-2883 | Sequence 2883, Ap |
| C 30 | 43.4 | 5.1 | 1356 | 4 US-09-307-106-49 | Sequence 49, Appl |
| C 31 | 42.8 | 5.0 | 705 | 4 US-09-270-767-5061 | Sequence 5061, Ap |
| C 32 | 42.8 | 5.0 | 705 | 4 US-09-270-767-20343 | Sequence 20343, A |
| C 33 | 42.8 | 5.0 | 5152 | 4 US-10-204-708-47 | Sequence 47, Appl |
| C 34 | 42.6 | 5.0 | 5152 | 4 US-10-204-708-73 | Sequence 73, Appl |
| C 35 | 42.6 | 5.0 | 319608 | 4 US-09-539-333D-1 | Sequence 1, Appl |
| C 36 | 42.6 | 5.0 | 319608 | 4 US-09-679-409-1 | Sequence 1, Appl |
| C 37 | 42.4 | 5.0 | 1449 | 4 US-09-107-532A-2018 | Sequence 2018, Ap |
| C 38 | 42.2 | 5.0 | 4074 | 1 US-08-471-033-19 | Sequence 19, Appl |
| C 39 | 42.2 | 5.0 | 4074 | 2 US-08-471-044-19 | Sequence 19, Appl |
| C 40 | 42.2 | 5.0 | 4074 | 2 US-08-463-483A-19 | Sequence 19, Appl |
| C 41 | 42.2 | 5.0 | 4074 | 2 US-08-471-046A-19 | Sequence 19, Appl |
| C 42 | 42.2 | 5.0 | 4074 | 2 US-08-470-566B-19 | Sequence 19, Appl |
| C 43 | 42.2 | 5.0 | 4074 | 2 US-08-469-334-19 | Sequence 19, Appl |
| C 44 | 42.2 | 5.0 | 4074 | 3 US-09-300-529-19 | Sequence 19, Appl |
| C 45 | 42 | 4.9 | 98844 | 4 US-09-791-211-10 | Sequence 10, Appl |

ALIGNMENTS

RESULT 1
US-08-363-010-3/c
; Sequence 3, Application US/08363010
; Patent No. 5512484
; GENERAL INFORMATION:
; APPLICANT: Yamamoto, Mika
; APPLICANT: Oheda, Kenji
; TITLE OF INVENTION: CARROT 16 KD PROTEIN, GENE CODING FOR
; TITLE OF INVENTION: SAID PROTEIN AND PLASMID CONTAINING SAID GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Birch, Stewart, Kolasch & Birch, LLP
; STREET: 8110 Gatehouse Road, Suite 500 East
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22042

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/363,010
FILING DATE: 23-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 20-3628P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 739 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Daucus carota
STRAIN: Kuroda Gosun
US-08-363-010-3

Query Match 14.5%; Score 123.6; DB 1; Length 739;
Best Local Similarity 96.9%; Pred. No. 9.3e-22;

| | Matches | 126; | Conservative | 0; | Mismatches | 4; | Indels | 0; | Gaps | 0; |
|----|---------|--|--------------|----|------------|----|--------|----|------|----|
| Qy | 722 | TGAGAAATGAATATGATTATTGGTTTATAAATAGTGTAATTTTATTTAAAAAGATCGCATAC | 781 | | | | | | | |
| Dd | 718 | TGAGAAATGAATATGATTATTGGTTTATAAATAGTGCAATTTTATTTAAAAGATCGCATAC | 659 | | | | | | | |
| Qy | 782 | ATTACCAGCCAGATGAAGTTATTTCATCACAACCTCACAAAGTACAAAGAAAAAGTTGC | 841 | | | | | | | |
| Dd | 658 | ATTACCAGCCAGATGAAGTTATTTCATCACAAGTCAACAAGTACAAAGAAAAGTTGC | 599 | | | | | | | |
| Qy | 842 | AATTCTGTCA | 851 | | | | | | | |
| Dd | 598 | AATTCTGTCA | 589 | | | | | | | |

```

1  RESULT 2
2  US-08-232-463-14/c
3  ; Sequence 14, Application US/08232463
4  ; Patent No. 5670367
5  ; GENERAL INFORMATION:
6  ; APPLICANT: DORNER, F.
7  ; APPLICANT: SCHEIFLINGER, F.
8  ; APPLICANT: FALKNER, F. G.
9  ; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
10 ; NUMBER OF SEQUENCES: 52
11 ; CORRESPONDENCE ADDRESS:
12 ; ADDRESSEE: Foley & Lardner
13 ; STREET: 1800 Diagonal Road, Suite 500
14 ; CITY: Alexandria
15 ; STATE: VA
16 ; COUNTRY: USA
17 ; ZIP: 22313-0299
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: Patent In Release #1.0, Version #1.25
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER:
25 ; US/08/232.463

```

| | Query Match | 9.1% | Score 77.8; | DB 1; | Length 7218; |
|----|-----------------------|--|--------------------|-----------|--------------|
| | Best Local Similarity | 9.6%; | Pred. No. 5.9e-10; | | |
| | Matches 40; | Conservative 219; | Mismatches 156; | Indels 0; | Gaps 0; |
| Qy | 155 | ATTATTGAGAAATACAACTTTGCAGAAAAGCCAACTCTCGGTTTTCAGGTTCTCGAATT | 214 | | |
| Db | 1521 | ATCTATTTCAGTTTCAAAAACGGCATGTAGGCATCACTGTAATTAACCTATCTATGCAAGT | 1462 | | |
| Qy | 215 | AATTAAAGCTTTTACATCACTATGGAAGTGATACATTATTCGAAGCCGAGTGGAGGGAAGAACT | 274 | | |

| Db | 1461 | AGTTAAAGAGATAGAGAATAATTTGGTACR | 1402 |
|--|------|--|------|
| Qy | 275 | TGAGTACGTTGAATGACAAAAATTCAGCAACACAGTACACAGAAGCTCTCGAAGTAATGTT | 334 |
| Db | 1401 | RR | 1342 |
| Qy | 335 | GAGAGTATGCAATAAGCAGAAATACGGGATACGTCACAAAGTATTTTACGACACAATAGA | 394 |
| Db | 1341 | RR | 1282 |
| Qy | 395 | AAAGCAGACAAAGGATAAAGTGCTCTATGAATGGTGAGGAAATGGGACGGTTTAGGTTAG | 454 |
| Db | 1281 | RR | 1222 |
| Qy | 455 | CTTAAAAAAGTGACCTTCTTGAAGTAAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT | 514 |
| Db | 1221 | RR | 1162 |
| Qy | 515 | AATAAATATAAGTTATTAAGTCTTTGGAAAAAATAAGAAGTTGTAAGAAAAAG | 569 |
| Db | 1161 | RR | 1107 |
| RESULT 3 | | | |
| US-08-911-434A-2/c | | | |
| : Sequence 2, Application US/08911434A | | | |
| : Patent No. 5959176 | | | |
| : GENERAL INFORMATION: | | | |
| : APPLICANT: TORIKAI, Satomi | | | |
| : APPLICANT: OEDA, Kenji | | | |
| : TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF | | | |
| : NUMBER OF SEQUENCES: 14 | | | |
| : CORRESPONDENCE ADDRESSES: | | | |
| : ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP | | | |
| : STREET: P.O. BOX 747 | | | |
| : CITY: FALLS CHURCH | | | |
| : STATE: VIRGINIA | | | |
| : COUNTRY: UNITED STATES OF AMERICA | | | |
| : ZIP: 22040 | | | |
| : COMPUTER READABLE FORM: | | | |
| : MEDIUM TYPE: Floppy disk | | | |
| : COMPUTER: IBM PC compatible | | | |
| : OPERATING SYSTEM: PC-DOS/MS-DOS | | | |
| : SOFTWARE: PatentIn Release #1.0, Version #1.30 | | | |
| : CURRENT APPLICATION DATA: | | | |
| : APPLICATION NUMBER: US/08/911,434A | | | |
| : FILING DATE: 12-AUG-1997 | | | |
| : CLASSIFICATION: 800 | | | |
| : ATTORNEY/AGENT INFORMATION: | | | |
| : NAME: Stewart, Raymond C. | | | |
| : REGISTRATION NUMBER: 21,066 | | | |
| : REFERENCE/DOCKET NUMBER: 2185-0199P | | | |
| : TELECOMMUNICATION INFORMATION: | | | |
| : TELEPHONE: (703)205-8000 | | | |
| : TELEFAX: (703)205-8050 | | | |
| : INFORMATION FOR SEQ ID NO: 2: | | | |
| : SEQUENCE CHARACTERISTICS: | | | |
| : LENGTH: 2042 base pairs | | | |
| : TYPE: nucleic acid | | | |
| : STRANDEDNESS: single | | | |
| : TOPOLOGY: linear | | | |
| : MOLECULE TYPE: DNA (genomic) | | | |
| : HYPOTHETICAL: NO | | | |
| : ORIGINAL SOURCE: | | | |
| : ORGANISM: Daucus carota L. | | | |
| : INDIVIDUAL ISOLATE: Kuroda Gosun | | | |
| : FEATURE: | | | |
| : NAME/KEY: promoter | | | |
| : LOCATION: 1..2042 | | | |
| US-08-911-434A-2 | | | |
| Query Match | | | |
| 8.3% | | | |
| Score 71: | | | |
| DB 2: | | | |
| Length 2042: | | | |

Best Local Similarity 73.6%; Pred. No. 1.9e-08;
Matches 117; Conservative 0; Mismatches 40; Indels 2; Gaps 2;

QY 428 GTGAGAAATGGACGGTTAGTTAGCTTAAAGAAAGTACTTCTTACTTGAAGTAAG 487

DB 1159 GGGATGAGAAGGCTGTTGGTTAAACTTAAAGAAAGTGC-TCTTGCATAAGTAAAG 1101

QY 488 AAGTGGAGTAGAAGTATGATAGTAAAGTAAATTAATAAGTTATAAGTGTGTTGGAAAG 547

DB 1100 AAGTGACACAGAAGTAAAGTAAATTAAGATTATAAGTATTAATTTGTTGGAAAG 1041

QY 548 AATAGCAAGTTGT-AAAGAAAGTATGACATTTTCTACTT 585

DB 1040 AAGCAGAAGCTGCGTAAAGTAAAGTTAGTATTGTCGGTTT 1002

RESULT 4

US-08-911-434A-5

; Sequence 5, Application US/08911434A

; Patent No. 5959176

; GENERAL INFORMATION:

; APPLICANT: TORIKAI, Satomi

; APPLICANT: OEDA, Kenji

; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP

; STREET: P.O. BOX 747

; CITY: FALLS CHURCH

; STATE: VIRGINIA

; COUNTRY: UNITED STATES OF AMERICA

; ZIP: 22040

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/911,434A

; FILING DATE: 12-AUG-1997

; CLASSIFICATION: 800

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Raymond C.

; REGISTRATION NUMBER: 21,066

; REFERENCE/DOCKET NUMBER: 2185-0199P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703)205-8000

; TELEFAX: (703)205-8050

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 836 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Daucus carota L.

; INDIVIDUAL ISOLATE: Kuroda Gosun

; FEATURE:

; NAME/KEY: terminator

; LOCATION: 1..836

US-08-911-434A-5

Query Match

Best Local Similarity 71.2%; Pred. No. 9.8e-07;

Matches 84; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 437 TGGACGGTTAGTTAGCTTAAAGAAAGTACTTCTTACTTGAAGTAAGTGGAGT 496

DB 518 TGGGGCAGTTTGGCTGGACTTAAAGAAAGTACTTATGCTTAAATAAATAGTAGAT 577

QY 497 AGAACTGATAAGTAAAGTAATAATTATAAGTTATAAGTGTGTTGGAAAGAAATAGA 554

DB 578 ATAAGTGAAAGTGTGATTGGACTTATAAGTTATAAGTGTGTTGAATATATATGA 635

RESULT 5

US-09-806-708B-22

; Sequence 22, Application US/09806708B

; Patent No. 6784342

; GENERAL INFORMATION:

; APPLICANT: The University of British Columbia

; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants

; FILE REFERENCE: 4810-58741

; CURRENT APPLICATION NUMBER: US/09/806,708B

; PRIOR FILING DATE: 2001-04-03

; PRIOR APPLICATION NUMBER: US 60/147,133

; PRIOR FILING DATE: 1999-08-04

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 22

; LENGTH: 1141

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (1)..(1141)

; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. PAF1 promoters

US-09-806-708B-22

Query Match

Best Local Similarity 6.6%; Score 55.8; DB 4; Length 1141;

Matches 48; Conservative 272; Mismatches 250; Indels 8; Gaps 1;

QY 220 AGCTTTTACATCACTATGGAAGTGATATCTTGAAGCGGAGTGGAGGAAGACTTGAGT 279

DB 57 RKWTWARMYCKRYRWNNKSRWKGWYKKYTCANNTSBRTHARKWDMKTYAYMTMN 116

QY 280 ACGTGAATGAGAAATTCAGCAACACAGTACAGAAAGCTCTGGAAGTAATGTTGAGAG 339

DB 117 KWKGTGRHRYWRWAMBTDVHHVYTMNNAWTTCMMDKDDKRTRWKKNNATGMD 176

QY 340 TATGCAATAAGGAATACGGGATACGTCAGCAAGTATTTTACGACACATAAGAAAGG 399

DB 177 DDTKYHNNNNGCBVTVMVRYKTRDWSBKRMNYGMBWQWNSYDVTYYWYWDMDCKR 236

QY 400 CAGAAAGGATAAGCTCTATGAATGCTGAGGAATTCGGACGGTTTAGGTTAGCTTAA 459

DB 237 KVRWVTRGRNRYNVAWBTARRRYNNNGWTBAMAYRWTTNNNNNNNAKAMCKRKYNG 296

QY 460 AAAAAAGTGACTTCTTAAAGTAATGAAGTGGAGTAGAACTGATAAGTAAA----- 512

DB 297 WNRABVNSTCTTWKSKTTKVRTSCWANNCRAGDANKDKHKWKSAAAGVYNNNNNNNWT 356

QY 513 -GTAATAATAAGTTATAAGTGTGTTGGAAAGAAATAGAGTGTGTAAGAAAGTT 571

DB 357 YKKARHBAARDWVWHSAMKWHANAHAHYSRKKWTBYRKRTVMNNNGTTTMMKEMWAWYK 416

QY 572 AGCATTTTCTACTTCCAACTTATTTCTCAGCACTTCTTAAAGTACTTCTTACTTTTAA 631

DB 417 MDMDWBTNNNNNGRTTYIGTKKKMMYYKKANNCKRWDHKTCHNNNTTWWKK 476

QY 632 CACAAACGGGTCAAGAAAGTGAAGCAAAAGCTGGAGTTACTTCTTATAAGAAATGTTT 691

DB 477 TYWNNCYKWSMTNGKSHREAAAVYTWYWWRRYAHANNNNNDWYWKACTYKYVCWSKW 536

QY 692 ATACTAAATGAGAAATGACAAACACAGAAATGAGAAATGAGTAATGATTATGTTTATAAA 751

DB 537 WNNYAAWYTKSSWNTYSRYRWKTNNSWRWSRSDTRSGMRANNYARABHYGYKWNTRWBW 596

QY 752 TAGTGTATTTTTAAAAAGATCGCATACATTACCAG 789

DB 597 SHTWBHBRAAAHYWMBMYBAKCHCMKAWYKAKKYAG 634

Db 177 HHWCATNNNNWWWYAYMHMHKKGKAAWNNKTAERDDHBAHVKTWYWRDYWCAMCW 118
QY 626 TTTTACACAAACGGCTCAAGAAAGTGAAGCABA 661
Db 117 MNAKAVRTAMRGMWYITDRYVSANNTGVRMMRW 82

RESULT 8

US-08-960-780-51
; Sequence 51, Application US/08960780
; Patent No. 6204435
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; TITLE OF INVENTION: No. 6204435el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 134
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/960,780
; FILING DATE: 30-OCT-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: MA-708
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 177C8
US-08-960-780-51

Query Match 5.3%; Score 44.8; DB 3; Length 1341;
Best Local Similarity 47.7%; Pred. No. 0.058;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;
QY 269 AGAACCTTGAGTGTGATGAGAAAATTCAGCAACACAGTCAAGAAAGCTCTGGAAGT 328
Db 67 ATATCTTTTAAATGAATGAAGTGAAGTGAACAAATTAATAATCTCAAGT 126
QY 329 AATGTTGAGAGTATGAATAAGAGAGATACGGGGATACGTCAAAGTATTTTTTACGACAC 388
Db 127 AATATATACTAATCTGCAAAATCTAAAATCACTGACAGGTAGAGGATTTTAAAGAAGAT 186

QY 389 AATAGAAAAGGCAGAAAGGATAAAGTGTCTTATGAATGGTGAGAAATTTGGGACGGTTTA 448
Db 187 AAGGAAAAGCCGAAAGAAATGGGGGAAAGAAAAGAGAGTGGAACTTAACCTGCTACT 246
QY 449 GGTTAGCTTTAAAAAAGTGAAGTCTTCTTACTTGAAGTAAAGAAAGTGGAGTGAAGTGAAG 508
Db 247 GAAAAAGGAAAAAATGAATAATTTTTTTAGATAATAAAAAATGATATAAAGACAAATTTATAA 306
QY 509 TAAAGTAAATAATATAAGTTTATAAAGTGTTTTGGAAAAAGAAATAGAAAGTTGTAAACAAAA 568
Db 307 GAAATTAATCTTTTCTATGGCAGGCTCA--TTTGAAGATGAATAAAGATTTTAAAGAAA 364
QY 569 GTTAGCATTCTTCTACTTCCAACTTATTTCTCAGCACTTCTTA 610
Db 365 TTGATAAGATGTTTGATAAAACCAATCTATCAAAATCTTATTA 406

RESULT 9

US-09-073-898-51
; Sequence 51, Application US/09073898
; Patent No. 6242669
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6242669el Pesticidal Toxins and Nucleotide
; TITLE OF INVENTION: Sequences Which Encode These Toxins
; NUMBER OF SEQUENCES: 144
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/073,898
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-708C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

```
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8
; US-09-073-898-51

Query Match      5.3%; Score 44.8; DB 3; Length 1341;
Best Local Similarity 47.7%; Pred. No. 0.058;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

Qy 269 AGAAGCTTGAGTACGTAATGAGAAATTCAGCAACACAGTACAGAAAGCTCTGGAAGT 328
Db 67 ATATCTTTTAAATTAATGAAGTGATAAAAGCTGAACCAATTAATTAATTAATTTCTCAAGT 126

Qy 329 AATGTTGAGAGTATCAATAAGAGAGATACGGGATAGCTCAAGATTTTTCAGCAC 388
Db 127 AATATATCTACTTGCATAAATCTAAATCACTGACAGGTAGAGATTTTAAAGAAGAT 186

Qy 389 AATAGAAAGGCGAAGGATTAAGTGTCTTATGAATGGTGAGGAATGGGACGGTTTA 448
Db 187 AAGGAAAGGCGAAGGATGGGGGAAAGAAAGAAAGAGTGGAACCTAATCTACT 246

Qy 449 GGTAGCTTAAAGAAAGTACTTCTTACTTGAAGTAAATGAAGTGGAGTAGAACTGATAAG 508
Db 247 GAAAGGAAAGAAATGAATATTTTATAGATATTAATATATATATATATATATAA 306

Qy 509 TAAAGTAAATATATATAGTTATTAAGTGTGTTTGGAAAGAAATAGAGTTCTTAAAGAAA 568
Db 307 GAAATTAATCTTTTCTATGGCAGGCTCA--TTTGAAGATGAATAAAGATTTTAAAGAAA 364

Qy 569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
Db 365 TTGATAAGATGTTTGATATAAACCAATCTATCAAAATTTCTATT 406

RESULT 10
US-09-371-913A-2
; Sequence 11, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Finstad Lee, Stacey
; APPLICANT: Walz, Mikki
; APPLICANT: Sturgis, Blake
; TITLE OF INVENTION: Pesticidal Toxins and Genes from Bacillus laterosporus
; FILE REFERENCE: MA-719XC2 US
; CURRENT APPLICATION NUMBER: US/09/371,913A
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/095,955
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/138,251
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Bacillus laterosporus
US-09-371-913A-2

Query Match      5.3%; Score 44.8; DB 3; Length 1341;
Best Local Similarity 47.7%; Pred. No. 0.058;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

Qy 269 AGAAGCTTGAGTACGTAATGAGAAATTCAGCAACACAGTACAGAAAGCTCTGGAAGT 328
Db 67 ATATCTTTTAAATTAATGAAGTGATAAAAGCTGAACCAATTAATTAATTAATTTCTCAAGT 126

Qy 329 AATGTTGAGAGTATCAATAAGAGAGATACGGGATAGCTCAAGATTTTTCAGCAC 388
Db 127 AATATATCTACTTGCATAAATCTAAATCACTGACAGGTAGAGATTTTAAAGAAGAT 186

Qy 389 AATAGAAAGGCGAAGGATTAAGTGTCTTATGAATGGTGAGGAATGGGACGGTTTA 448
Db 187 AAGGAAAGGCGAAGGATGGGGGAAAGAAAGAAAGAGTGGAACCTAATCTACT 246

Qy 449 GGTAGCTTAAAGAAAGTACTTCTTACTTGAAGTAAATGAAGTGGAGTAGAACTGATAAG 508
Db 247 GAAAGGAAAGAAATGAATATTTTATAGATATTAATATATATATATATATAA 306

Qy 509 TAAAGTAAATATATATAGTTATTAAGTGTGTTTGGAAAGAAATAGAGTTCTTAAAGAAA 568
Db 307 GAAATTAATCTTTTCTATGGCAGGCTCA--TTTGAAGATGAATAAAGATTTTAAAGAAA 364

Qy 569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
Db 365 TTGATAAGATGTTTGATATAAACCAATCTATCAAAATTTCTATT 406

US-09-371-913A-2
; Sequence 2, Application US/09371913A
; Patent No. 6297369
; GENERAL INFORMATION:
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Finstad Lee, Stacey
; APPLICANT: Walz, Mikki
; APPLICANT: Sturgis, Blake
; TITLE OF INVENTION: Pesticidal Toxins and Genes from Bacillus laterosporus
; FILE REFERENCE: MA-719XC2 US
; CURRENT APPLICATION NUMBER: US/09/371,913A
; CURRENT FILING DATE: 1999-08-10
; PRIOR APPLICATION NUMBER: 60/095,955
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/138,251
; PRIOR FILING DATE: 1999-06-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1341
; TYPE: DNA
; ORGANISM: Bacillus laterosporus
US-09-371-913A-2

Query Match      5.3%; Score 44.8; DB 3; Length 1341;
Best Local Similarity 47.7%; Pred. No. 0.058;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

Qy 269 AGAAGCTTGAGTACGTAATGAGAAATTCAGCAACACAGTACAGAAAGCTCTGGAAGT 328
Db 67 ATATCTTTTAAATTAATGAAGTGATAAAAGCTGAACCAATTAATTAATTAATTTCTCAAGT 126

Qy 329 AATGTTGAGAGTATCAATAAGAGAGATACGGGATAGCTCAAGATTTTTCAGCAC 388
Db 127 AATATATCTACTTGCATAAATCTAAATCACTGACAGGTAGAGATTTTAAAGAAGAT 186

Qy 389 AATAGAAAGGCGAAGGATTAAGTGTCTTATGAATGGTGAGGAATGGGACGGTTTA 448
Db 187 AAGGAAAGGCGAAGGATGGGGGAAAGAAAGAAAGAGTGGAACCTAATCTACT 246

Qy 449 GGTAGCTTAAAGAAAGTACTTCTTACTTGAAGTAAATGAAGTGGAGTAGAACTGATAAG 508
Db 247 GAAAGGAAAGAAATGAATATTTTATAGATATTAATATATATATATATATAA 306

Qy 509 TAAAGTAAATATATATAGTTATTAAGTGTGTTTGGAAAGAAATAGAGTTCTTAAAGAAA 568
Db 307 GAAATTAATCTTTTCTATGGCAGGCTCA--TTTGAAGATGAATAAAGATTTTAAAGAAA 364

Qy 569 GTTAGCATTTTCTACTTCCAACTTATTTCTCACGACTTCTTA 610
Db 365 TTGATAAGATGTTTGATATAAACCAATCTATCAAAATTTCTATT 406

US-09-307-106-11
; Sequence 11, Application US/09307106
; Patent No. 6603063
; GENERAL INFORMATION:
; APPLICANT: Feitelson, Jerald S.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Stockhoff, Brian A.
; APPLICANT: Schmeits, James
; APPLICANT: Loewer, David
; APPLICANT: Dullum, Charles Joseph
; APPLICANT: Muller-Cohn, Judy
; APPLICANT: Stamp, Lisa
; APPLICANT: Morrill, George
; APPLICANT: Finstad-Lee, Stacey
; TITLE OF INVENTION: No. 6603063el Pesticidal Toxins and Nucleotide
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/307,106
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/029,848
; FILING DATE: 30-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/960,780
; FILING DATE: 30-OCT-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/073,898
; FILING DATE: 05-MAY-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-708C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1341 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
```



```
;
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: PS177C8
; SEQUENCE DESCRIPTION: SEQ ID NO: 51:
US-09-850-351A-51

Query Match          5.3%; Score 44.8; DB 4; Length 1341;
Best Local Similarity 47.7%; Pred. No. 0.058;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

QY 269 AGAAGCTTGGAGTATGAGTAAAGGAGATACGGGGATACGCTCAAAAGTATTTTACGACAC 328
Db 67 ATATCTTTTAAATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 126
QY 329 AATGTTGAGAGTATGAGTAAAGGAGATACGGGGATACGCTCAAAAGTATTTTACGACAC 388
Db 127 AATATATACTAATTCGAAATCTAAATATCACTGACAGGTAGAGGATTTTAAAGAAGAT 186
QY 389 AATAGAAAGGAGAGAGGATTAAGTCTCTATGAATGGTGAGGAATTTGGGACCGTTTA 448
Db 187 AAGGAAAAAGCGAAAGATGCGGGGAGAAAGAAAGAAAGAGTGGAAACTAACTGCTACT 246
QY 449 GGTAGCTTAAAAAAGTACTTCTTACTTGAAGTAAATGAAGTGAAGTGAAGTGAAGTGAAG 508
Db 247 GAAAGAGGAAATGAATATTTTATAGATAATAAAATGATATAAGACAAATTTATATA 306
QY 509 TAAAGTAAATAATATAAGTTATTAAGTGTGTTGGAAAAAGAAATAGAAAGTTGTAAGAAAA 568
Db 307 GAAATTAATCTTTTCTATGCGAGCTCA--TTTGAAGATGAATATAAGATTTTAAAGAAA 364
QY 569 GTTAGCATTTTCTACTTCCAACTTATTTCTCAGGACTTCTTA 610
Db 365 TTGATAAGATGTTTGATAAAACCAATCTATCAAAATCTTATTA 406
```

RESULT 14

```
US-08-471-033-22
; Sequence 22, Application US/08471033
; Patent No. 5770696
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; ADDRESS: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30B
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471,033
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/314,594
; FILING DATE: 09-SEP-1994
; PRIOR APPLICATION DATA:
```

```
;
; APPLICATION NUMBER: US 08/218,018
; FILING DATE: 23-MAR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/037,057
; FILING DATE: 25-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Pace, Gary M.
; REGISTRATION NUMBER: P-40,403
; REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV7 - SOLV3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8582
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4041 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4038
; OTHER INFORMATION: /product= "VIPIA(a)/VIP2A(a) fusion
; OTHER INFORMATION: product"
US-08-471-033-22

Query Match          5.3%; Score 44.8; DB 1; Length 4041;
Best Local Similarity 47.7%; Pred. No. 0.086;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;

QY 269 AGAAGCTTGGAGTATGAGTAAAGGAGATACGGGGATACGCTCAAAAGTATTTTACGACAC 328
Db 88 ATATCTTTTAAATAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 147
QY 329 AATGTTGAGAGTATGAGTAAAGGAGATACGGGGATACGCTCAAAAGTATTTTACGACAC 388
Db 148 AATATATACTAATTCGAAATCTAAATATCACTGACAGGTAGAGGATTTTAAAGAAGAT 207
QY 389 AATAGAAAGGAGAGAGGATTAAGTCTCTATGAATGGTGAGGAATTTGGGACCGTTTA 448
Db 208 AAGGAAAAAGCGAAAGATGCGGGGAGAAAGAAAGAGTGGAAACTAACTGCTACT 267
QY 449 GGTAGCTTAAAAAAGTACTTCTTACTTGAAGTAAATGAAGTGAAGTGAAGTGAAGTGAAG 508
Db 268 GAAAGAGGAAATGAATATTTTATAGATAATAAAATGATATAAGACAAATTTATATA 327
QY 509 TAAAGTAAATAATATAAGTTATTAAGTGTGTTGGAAAAAGAAATAGAAAGTTGTAAGAAAA 568
Db 328 GAATTAATCTTTTCTATGCGAGCTCA--TTTGAAGATGAATATAAGATTTTAAAGAAA 385
QY 569 GTTAGCATTTTCTACTTCCAACTTATTTCTCAGGACTTCTTA 610
Db 386 TTGATAAGATGTTTGATAAAACCAATCTATCAAAATCTTATTA 427
```

RESULT 15

```
US-08-471-044-22
; Sequence 22, Application US/08471044
; Patent No. 5840868
; GENERAL INFORMATION:
; APPLICANT: Warren, Gregory W
; APPLICANT: Koziel, Michael G
; APPLICANT: Mullins, Martha A
; APPLICANT: Nye, Gordon J
; APPLICANT: Carr, Brian
; APPLICANT: Desai, Nalini M
; APPLICANT: Kostichka, N. Kristy
; APPLICANT: Duck, Nicholas B
; APPLICANT: Estruch, Juan J
; TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
```


STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30B
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,044
FILING DATE: 06-JUN-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/463,483
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Pace, Gary M.
REGISTRATION NUMBER: 40,403
REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SQLv3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8582
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 4041 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..4038
OTHER INFORMATION: /product= "VIP1A(a)/VIP2A(a) fusion"
OTHER INFORMATION: product"
US-08-471-044-22

Query Match 5.3%; Score 44.8; DB 2; Length 4041;
Best Local Similarity 47.7%; Pred. No. 0.086;
Matches 163; Conservative 0; Mismatches 177; Indels 2; Gaps 1;
QY 269 AGAAGCTTGAGTAGTGGAATGAGAAAAATTCAGCAACACACAGTACAGAAAGCTCTGGAAGT 328
DB 88 ATATCTTTTATTAATTAATGAAGTGATAAAGCTGAACAATTAATATAATTAATCTCAAGT 147
QY 329 AATGTTGAGAGTATGCAATAGGAGAAATACGGGGATACGCTCAAGTAATTTTTTACGACAC 388
DB 148 AAATATACTAAGCTTGCAAAATCTAAAAATCACTGACAAAGTAGAGGATTTTAAAGAGAT 207
QY 389 AATAGAAAGGAGAAAGGATAAGTGCTCTATGATGCTGAGGAATTTGGGACGGTTTA 448
DB 208 AAGGAAAAAGCGAAAGAAATGGGGGAAAGAAAAAGAGAGTGGAAACTAACTGCTACT 267
QY 449 GGTAGCTTTAAAAAGTCACTTCTTACTTTGAAGTAATGAAGTGGAGTAGTAAGTATAG 508
DB 268 GAAAGAGGAAATGAATTAATTTTTTATTAATTAATTAATGAATTAAGCAATTAATAA 327
QY 509 TAAAGTAATTAATTAAGTTATTAAGTGTGTTGAAAGAAATAGAAAGTTGTAAGAAAA 568
DB 328 GAAATTACTTTTTCTATGCGAGCTCA--TTTGAAGATGAATATAAGATTTTAAAGAAA 385
QY 569 GTTAGCAATTTCTACTTCCAACTTAATTTCTCAGGACTTCTTA 610

Db 386 TTGATAAGATGTTTGATATAAACCAATCTATCAAAATCTTATTA 427
Search completed: December 6, 2004, 22:52:27
Job time : 86.3031 secs

THIS PAGE BLANK (USPTO)


```
Db      84138 TTATCTAATCTATATATCAAACTACCTCAAAATTTAATACTTAATAAACAATATCT 84079
Qy      636 AACGGTCAAGGAAGTGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATAC 695
Db      84078 ATTAACCTCAAAAAATCTACACACACAAATAAAATAAATTC-TATAAAATAACAAAA 84020
Qy      696 TAAATGAGAAATGACAAACACAGAAATGGAATGAATATGATTATGTTTAAATATAGT 755
Db      84019 TTAATAAATAAAAAAATTTTACAATTAATAAATAAATAATTTATTAATACAAATAAAAA 83960
Qy      756 GTATTTTATTTAAAAAGATCGCATACATTACCAGCCAGATGAAGTTATTTCATCAACATC 815
Db      83959 AATTTTATCCAAATCCTTCAAAAAACAATACCAAAATAAATCTTAAAAATCTAGCAT 83900
Qy      816 ACAACAAA 823
Db      83899 CTTATAAA 83892
```

RESULT 2

```
US-10-257-166-162/c
; Sequence 162, Application US/10257166
; Publication No. US20040023230A1
```

GENERAL INFORMATION:

```
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Method and Nucleic Acids for Analysing the Methylation of
; TITLE OF INVENTION: Genes Implicated in Pharmacogenomics
; FILE REFERENCE: 5013.1011
; CURRENT APPLICATION NUMBER: US/10/257,166
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: PCT/EP01/07470
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-06-29
; 2000-06-30
; NUMBER OF SEQ ID NOS: 178
; SEQ ID NO 162
; LENGTH: 2114
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-257-166-162
```

```
Query Match      6.4%; Score 54.6; DB 16; Length 2114;
Best Local Similarity 56.4%; Pred. No. 0.068;
Matches 102; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy      649 AAGTGGAGCAAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTACTAAATGAGAAATG 708
Db      1499 AATAAATAACCAAAAAATATTTTAATCATATATTATATATATATAATAAATAAACA 1440
Qy      709 ACAACACAGAAATGAGAAATGAATATGATTGTGTTTAAATAATAGTGTATTTTAA 768
Db      1439 ACAATAATACAAAAAATAAAATAAATAAATAAATAAATAATCTATTATTATTATTTAA 1380
Qy      769 AAAGATCCGATACATTACCAGCCAGATGAAGTTATTCATCAACTCAGCAAGATGAC 828
Db      1379 AATAAAATCTTACTTTACTTACCCAAATTAATAAATAAATAAATAATCTCAACTCTACA 1320
Qy      829 A 829
Db      1319 A 1319
```

RESULT 3

```
US-10-240-589C-130
; Sequence 130, Application US/10240589C
; Publication No. US20040076956A1
```

```
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: DNA repair
; FILE REFERENCE: 5013.1008
; CURRENT APPLICATION NUMBER: US/10/240,589C
; CURRENT FILING DATE: 2003-09-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03972
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 148
; SEQ ID NO 130
; LENGTH: 4045
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-589C-130
```

```
Query Match      6.2%; Score 52.8; DB 16; Length 4045;
Best Local Similarity 47.5%; Pred. No. 0.22;
Matches 188; Conservative 0; Mismatches 207; Indels 1; Gaps 1;

Qy      374 TATTTTTTACGACACAATAGAAAAGCGCAGAAAGGATAAAGTGCTCTATGAATGGTGAGG 433
Db      858 TTTGTTTTTAAAGAAAATATATTTTGTAAAAAAGAAAAAGAAAAAGAAAAAGAAAAA 917
Qy      434 AATTGGGACGGTTTAGGTTAGCTTAAAAAAGTGCCTCTTACTTGAAGTAAAGTGG 493
Db      918 AAGGGGGGAGTGTAGTGGAAATAGAGTATATAAATGTTTAAGTGTTTTAAATAAATAA 977
Qy      494 AGTAGAAGTAAAGTAAAGTAAATATATAAGTTATTAAAAAGTGTTCGAAAAAGAAATAG 553
Db      978 AATAAAGCTAAATGTTTGGGGGAGAGTTTTTTTTTAAAGTATTTCTTAAAGTTAGAG 1037
Qy      554 AAGTGTGTAAGAAAAGTTAGCATTTTCTACTTCCAACTTATTTCTCAGACTTCTTAAAA 613
Db      1038 AAGAAATAATGGGTATTTTGTAAAGTTAGGTAAGATATTTTTTAAAAATTTTATAAGTG 1097
Qy      614 GTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAG-CAAAAAGCTGGAGTT 672
Db      1098 TAAATTTTTTTAGTTTTTAAATATTTATATAAGTTTTTATTGGAAGATAAATATATTAGAGT 1157
Qy      673 ACTTCTTATAAGAATGTTTATATACTAAATGAGAAATGACAAAACAGAGAAATGAGAAATGAAT 732
Db      1158 AATTTAATATGGTTTAAATTTATGTAATGTTAAATAATTTTTTTAAAAAAGATTTTAAAG 1217
Qy      733 ATGATTATGTTTAAATAAATAGTGTATTTTATTTAA 768
Db      1218 TAAATGTATGGTTTTATGAATAGAAATTTATGTTTTTA 1253
```

RESULT 4

```
US-10-473-126-386
; Sequence 386, Application US/10473126
; Publication No. US20040234973A1
```

GENERAL INFORMATION:

```
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Methods and nucleic acids for the analysis of hematopoietic cell
; TITLE OF INVENTION: proliferative disorders
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/473,126
; CURRENT FILING DATE: 2003-09-26
; NUMBER OF SEQ ID NOS: 1258
```



```
QY 374 TATTTTACGACACAATAGAAAAGGCGAGAAAGGATAAAGTGTCTCTATGAATGGTGAGG 433
Db 415 TTTTNTCCCNNAAAAAAANNCGCCCAANAAAAAANANGCANAAACAATTTTTTTT 356
QY 434 AATTGGGACGGTTTAGGTTAGCTTAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGG 493
Db 355 TTTNGGGNANTTTTTTTTTNNCNTNCAAAAANAATNANATTTNANAATATTTNANAATN 296
QY 494 AGTAGAAGTGAAGTAAAGTAATATTAAGTTTATTAAGTTTATTAAGTGTGGAAGAAAGAAATAG 553
Db 295 TNNTGAAAAAANAANAATNTTTTTTTTCATCAAAAANNTTATCCNCCNNAAAAA 236
QY 554 AAGTGTGAAGAAAGTAAAGTAATATTAAGTTTATTAAGTGTGGAAGAAAGAAATAG 613
Db 235 TTTNGGGNANTTTTTTTTTNNCNTNCAAAAANAATNANATTTNANAATATTTNANAATN 176
QY 614 GTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGAAGCAAAAAGCTGGAGTTA 673
Db 175 AAAANNTTTTTTTTTTCAAAAANAANAATTAATAAANAANAANAANAANAANAANAANA 116
QY 674 CTTCCTTATAAGATGTTTATCTAATAATGAGAAATGACAAACACAGAAATGAGAATGAATA 733
Db 115 AAAAAAANAANAANAATTAATAAANAANAANAANAANAANAANAANAANAANAANAANA 56

RESULT 7
US-09-814-353-12133/c
; Sequence 12133, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-03-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12133
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
; LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
; LOCATION: 293, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
; LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
; LOCATION: 451, 455, 468, 469, 472, 474, 475, 478, 481, 488, 490,
; LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
; OTHER INFORMATION: n = A,T,C or G

US-09-814-353-12133
```

```
Query Match 6.1%; Score 52; DB 10; Length 539;
Best Local Similarity 41.1%; Pred. No. 0.15; Mismatches 0; Gaps 0;
Matches 148; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 374 TATTTTACGACACAATAGAAAAGGCGAGAAAGGATAAAGTGTCTCTATGAATGGTGAGG 433
Db 415 TTTTNTCCCNNAAAAAAANNCGCCCAANAAAAAANANGCANAAACAATTTTTTTT 356
QY 434 AATTGGGACGGTTTAGGTTAGCTTAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGG 493
Db 355 TTTNGGGNANTTTTTTTTTNNCNTNCAAAAANAATNANATTTNANAATATTTNANAATN 296
QY 494 AGTAGAAGTGAAGTAAAGTAATATTAAGTTTATTAAGTGTGGAAGAAAGAAATAG 553
Db 295 TNNTGAAAAAANAANAATNTTTTTTTTCATCAAAAANNTTATCCNCCNNAAAAA 236
QY 554 AAGTGTGAAGAAAGTAAAGTAATATTAAGTTTATTAAGTGTGGAAGAAAGAAATAG 613
Db 235 TTTNTNGGGNNTTAGTCNNTTTTTNACNTTTTTTGCNCCNTTTTNATATTTTTTTANA 176
QY 614 GTACTTCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGAAGCAAAAAGCTGGAGTTA 673
Db 175 AAAANNTTTTTTTTTTCAAAAANAANAATTAATAAANAANAANAANAANAANAANAANA 116
QY 674 CTTCCTTATAAGATGTTTATCTAATAATGAGAAATGACAAACACAGAAATGAGAATGAATA 733
Db 115 AAAAAAANAANAANAATTAATAAANAANAANAANAANAANAANAANAANAANAANAANA 56

RESULT 8
US-10-311-455-1504/c
; Sequence 1504, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1504
; LENGTH: 13712
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-311-455-1504
```

```
Query Match 6.1%; Score 52; DB 15; Length 13712;
Best Local Similarity 50.0%; Pred. No. 0.56;
Matches 157; Conservative 0; Mismatches 155; Indels 2; Gaps 1;

QY 516 ATAATATAAGTTTATTAAGTTTGGAAAAGAAATAGAAAGTTGTAAGAAAAGTTAGCA 575
Db 10091 AAACATAACAAAATACAAAATATATATAAAAAAACAATCATATAAAAAAACAACA 10032
QY 576 TTTTCTACTTCCCACTTATTTCTCAGCACTTCTTAAAGTAGTCTTCTTACTTTTACACA 635
Db 10031 TCTTCAATAACAATCTATTTTAAATAAAAAACAANAATCAATTAATTAAAAAATAATAC 9972
QY 636 AACGGGTCAAGGAAAGTGAAGCAAAAAGCTGGAGTTACTTCTTATAAGAAATGTTTATAC 695
Db 9971 TAATTAACAAAAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 9912
```



```
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(3322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; NUMBER OF SEQ ID NOS: 2003-04-28
; SEQ ID NO 647
; LENGTH: 802
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; LOCATION: (1)..(802)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRP4577_100591C.1
US-10-425-115-647

Query Match          5.9%; Score 50.6; DB 18; Length 802;
Best Local Similarity 44.9%; Pred. No. 0.35;
Matches 227; Conservative 0; Mismatches 277; Indels 2; Gaps 1;

QY 265 AGGAAGAAGCTTGAGTACGTCGTAATGAGAAAAATTCAGCAACACAGTACAGAAAGCTCTGG 324
DB 745 AGAAGAAGAAAAAGAAAAAATAATTTTAAAAATTAATATAAATAAATAAATAAATAAATAA 686

QY 325 AGTAAATGTTGAGATGATGCAATAGGAGAAATACGGGGATACGTCAGAAAGTATTTTACG 384
DB 685 AAAATAATATAAATAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 626

QY 385 ACACAATAGAAAAGGCAGAAAGGATAAAGTCTCTATGAATGGTGAGGAATTTGGACGG 444
DB 625 AAAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 566

QY 445 TTTAGGTAGCTTAAAAAAGTGACTTCTTCTGAAAGTAATGAAAGTGGAGTAGAACTGA 504
DB 565 AAAATAATAATAAATAAATAAGACAAATATATAGAAAGTAATAAATAAATAAAGAAAGAA 506

QY 505 TAAAGTAAAGTAATAATTAAGTATTAAGTGTGGAAAGAAATAGAAAGTGTCTAAAG 564
DB 505 AAACAAAAAATAGAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAGTTATGCAA 446

QY 565 AAAAGTTAGCAATTTCTACTTCCAACTTATTTCTCAGCACTTCTTAAAAAGTACTTCTTAC 624
DB 445 TAAAGAAATAAATTTGTATGACAGACAAATATATTAACAACAATAAATAAAGAAATAT 386

QY 625 TTTTATACAAACGGGTCAAGGAAGTGGAGCAAAAAAGCTGGAGTTACTTCTTTATAAG 684
DB 385 AAAAAAATTAACAATAATAAGATAAATAA--AAATAATAATAATAAATAAATAAATAA 328

QY 685 AATGTTTACTAATAGCAATAGCAACACAGAAATGAGAAATGAGAAATGATGATTTGGT 744
DB 327 AACATAAAAACAATTAATAAATAATTCAAAAAACATTTAAAAAGTAACAAAAAAGTTAATATA 268

QY 745 TTAATAATAGTCTATTTTATTTAAAA 770
DB 267 TTAATAGAAGAAAAAAGTTAAAA 242

RESULT 12
US-10-311-455-1473/c
; Sequence 1473, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1473

Query Match          5.9%; Score 49.8; DB 15; Length 9157;
Best Local Similarity 45.7%; Pred. No. 1.4;
Matches 174; Conservative 0; Mismatches 207; Indels 0; Gaps 0;

QY 456 TTAATAAAGTGACTTCTTACTTGAAGTAATGAAGTCGAGTAGAAGTAAAGTAAAGTAA 515
DB 6328 TTTATAAATTTAAATTTTCAACGCTACATTTATTTACATAAAAAACTACTAACAAATCGTATA 6269

QY 516 ATAATATTAAAGTTATTAAGTGTGGAAAAAGAAATAGAAAGTTGTAAAGAAAAAGTTAGCA 575
DB 6268 ATTCGGTTTATATAAATATTTCAAAATTAACAAATATAAATACAAAAATAAATTAATA 6209

QY 576 TTTTCTACTTCCAACTTATTTCTCAGCACTTCTTAAAGTACTTCTTACTTTTACACA 635
DB 6208 ATTACCAAAAACTACAAATTTAAAAATAACAACTAAAAAAAATAAATAACGTTTTTAAAAA 6149

QY 636 AACGGGTCAAGGAAGTCGAGCAAAAGCTGGAGTACTTCTTATAAGATGTTTATAC 695
DB 6148 TATAAATAATTTTAAAAATTTTCTTAATAATAACTACAACTCTCAATAATCTTAAAAATC 6089

QY 696 TAAATGAGAAATGACAAACACAGAAATGAGAAATGAATATGATTTATTTGTTTAAATAAGT 755
DB 6088 CATTAATATATATACCTTAATAATAATAAATTTCTATATATATATATCTCAATAAAACTT 6029

QY 756 GTATTTTATTTAAAAAGATCGCATACATTACCGCAGATGGAAGTTATTCATCACACTC 815
DB 6028 TTTTTTTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5969

QY 816 ACAACAAGTACAAAGAAAA 836
DB 5968 AACAAAACTAAAAAATAAAAA 5948

RESULT 13
US-10-311-455-1701/c
; Sequence 1701, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311.455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1701
```



```
; SEQ ID NO 1701
; LENGTH: 17721
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1701

Query Match          5.8%; Score 49.6; DB 15; Length 17721;
Best Local Similarity 47.2%; Pred. No. 2.1; Indels 0; Gaps 0;
Matches 151; Conservative 0; Mismatches 169;

QY      510 AAGAGTAATAATTATATAAGTTTATTAAAGTGTTTCGGAAAAGAAAAATAGAAGTTGTAAAGAAAAAG 569
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      14890 AAAATAATTAAAAATAACAAATTTTTTTACTTCCAATAAAAAATAAACAACACTATATAAAA 14831

QY      570 TTAGCAATTTCTTACTCCAACTTATTTCTCAGACTTCTTAAAGTACTTCTTACTTTTTT 629
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      14830 AA'TTAATAAACCAAAT'CCAAATAAAATCCAA'TTATACTATATAAAATAAAAT'AAAA'TTAA 14771

QY      630 TACACAAACGGGTCGAAGGAAAGTGGAAAGCAAAAAGCTGGAGTACTCTCTTAAGAAGTGT 689
        ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      14770 TATACTTAAAAAATATTTTAAATTTAATATAATATTAATCAATAAATTCCAAAAAAATTTA 14711

QY      690 TTATACTAAATGAGAAATGACAAACACAGAAATGAGATGAATATGATATTATTTGGTTTAAT 749
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      14710 'TTATAAAATTCATAAA'TTATAATATTACTTTTATCATATTTTCC'TATTTTTTTTTTTTATTT 14651

QY      750 AATAGTGTATTTTATTTTAAAAAGATCGCATACATTACCAGGCCAGATGAAGTTTATTCATCA 809
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      14650 TATTTTATTTTCTCTTAAACCAAAATCTCGGTATACCACCACCAACTAAAATCAATAAC 14591

QY      810 CAACTCACAAACAAGTACAA 829
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      14590 ATAATCTCGACTCACTACAA 14571
        ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 14
US-10-425-115-160787/c
; Sequence 160787, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 160787
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(578)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MFT4577_78218C.1
; US-10-425-115-160787

```

| | | | | |
|-----------------------|-----------------|---|-----------|-------------|
| Query Match | 5.8% | Score 49.4; | DB 18; | Length 578; |
| Best Local Similarity | 53.2%; | Fold. No. 0.56; | | |
| Matches 124; | Conservative 0; | Mismatches 108; | Indels 1; | Gaps 1 |
| Qy | 480 | AAGTAATGAAGTGGAGTGAAGCTGATAAGTAAAGTAATATTAAGTTATTTAAAGTGT | 539 | |
| Db | 359 | AAATAAAGAAAGATGATATAAAATATAAAAAAAATATGAAATATAAAAAATATAAAGAA | 300 | |
| Qy | 540 | TGAAAAAGAAATAGAAGTTGTAAAGAAAGTTAGCATTTTCTACTTCCAACTTATTTCTC | 599 | |

| Db | Qy | Db | Qy | Db |
|-----------|-------------|-------------|-------------|-------------|
| RESULT 15 | US-10-311 | US-10-311 | US-10-311 | US-10-311 |
| | Sequence | Sequence | Sequence | Sequence |
| | Publication | Publication | Publication | Publication |
| | GENERAL | GENERAL | GENERAL | GENERAL |
| | APPLIC | APPLIC | APPLIC | APPLIC |
| | APPLIC | APPLIC | APPLIC | APPLIC |
| | APPLIC | APPLIC | APPLIC | APPLIC |
| | TITLE | TITLE | TITLE | TITLE |
| | FILE | FILE | FILE | FILE |
| | CURRENT | CURRENT | CURRENT | CURRENT |
| | PRIOR | PRIOR | PRIOR | PRIOR |
| | PRIOR | PRIOR | PRIOR | PRIOR |
| | PRIOR | PRIOR | PRIOR | PRIOR |
| | PRIOR | PRIOR | PRIOR | PRIOR |
| | PRIOR | PRIOR | PRIOR | PRIOR |
| | NUMBER | NUMBER | NUMBER | NUMBER |
| | SEQ ID | SEQ ID | SEQ ID | SEQ ID |
| | LENGTH | LENGTH | LENGTH | LENGTH |
| | TYPE: | TYPE: | TYPE: | TYPE: |
| | ORGAN | ORGAN | ORGAN | ORGAN |
| | FEATUR | FEATUR | FEATUR | FEATUR |
| | OTHER | OTHER | OTHER | OTHER |
| | US-10-311 | US-10-311 | US-10-311 | US-10-311 |

| Query Match | 5.8%; | Score 49.2; | DB 15; | Length 113515; |
|-----------------------|-----------------|---|-----------|----------------|
| Best Local Similarity | 46.7%; | Pred. No. 5.6; | | |
| Matches 156; | Conservative 0; | Mismatches 178; | Indels 0; | Gaps 0; |
| Qy | 499 | AACTGATAGTAAAGTAATAATTAAAGTTATTAAAGTTGTGGAAAAAGAAATAGAACTT | 558 | |
| Db | 13895 | AACACCAATTAATAATATACATATAAACAATCCCAATCTTACATATATTAAACAC | 13836 | |
| Qy | 559 | GTAAGAAAAAGTTAGCAATTTTCTACTCCCACTTATTCTCACGACTTCTTTAAAGTACT | 618 | |
| Db | 13835 | CACATACTTAATAATATAAACTACTCCTTTTCTTTTCTACATTTTTTTAAATCAA | 13776 | |
| Qy | 619 | TCTTACTTTTTTACACAAACGGGTCAAGGAAAGTGGAAAGCAAAAAGCTGGAGTTACTTCT | 678 | |
| Db | 13775 | CAAAAACAACTAATAATAATAATAAAATAAAATAATAAAAAATAAACTTTAAATTAATCAA | 13716 | |
| Qy | 679 | TATAGAATGTTTATCTAAATGCAAAATGACAAACACAGAAATGAGAAATGAATATGATT | 738 | |
| Db | 13715 | AACAATATTCAATTAACATAATTAATAATACAAAACACAAAAAAAACATCTTTTA | 13656 | |
| Qy | 739 | ATTGGTTTAATAATAGTGTATTTTATTAAAAAGATCGCATACATTACCAGCCAGATGAA | 798 | |
| Db | 13655 | ATTAAATCCTTAAATATATATTTCAATCAAATAAATAACAAAAAATAACACACAAAAA | 13596 | |
| Qy | 799 | GTTATTTCATCAACTCAACACAAAGTACAAAGA | 832 | |
| Db | 13595 | ACTAACATTTTACATTACCACCAATATAAAAAAAA | 13562 | |

Search completed: December 7, 2004, 06:11:21
Job time : 509.824 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:35:24 ; Search time 3211.92 Seconds
(without alignments)
9654.718 Million cell updates/sec

Title: US-09-806-197-2

Perfect score: 851

Sequence: 1 ctgaagaagaagtctcatoga.....aaaaagttgcaattctgtca 851

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:**

2: gb_est2:**

3: gb_hic:**

4: gb_est3:**

5: gb_est4:**

6: gb_est5:**

7: gb_est6:**

8: gb_gsa1:**

9: gb_gsa2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| C 1 | 83.4 | 9.8 | 1101 | 9 | AL063921 Drosophil |
| C 2 | 65.6 | 7.7 | 938 | 9 | CNS0067U |
| C 3 | 64 | 7.5 | 987 | 9 | CNS014FQ |
| C 4 | 63.8 | 7.5 | 1001 | 9 | CNS0155H |
| C 5 | 63.6 | 7.5 | 1092 | 9 | CNS020K7 |
| C 6 | 63.4 | 7.5 | 1101 | 9 | CNS00K95 |
| C 7 | 62.8 | 7.4 | 1201 | 9 | CNS0167M |
| C 8 | 62.6 | 7.4 | 1101 | 9 | CNS00KAE |
| C 9 | 61.8 | 7.3 | 1380 | 9 | AG321093 |
| C 10 | 61.2 | 7.2 | 1101 | 9 | CNS016L1 |
| C 11 | 60.8 | 7.1 | 1101 | 9 | CNS0039L |
| C 12 | 60.8 | 7.1 | 1101 | 9 | CNS0182P |
| C 13 | 60.6 | 7.1 | 1101 | 9 | CNS000B8 |
| C 14 | 60.2 | 7.1 | 1101 | 9 | CNS0039R |
| C 15 | 60.2 | 7.1 | 1101 | 9 | CNS00EVL |
| C 16 | 60 | 7.1 | 1146 | 9 | CNS021G2 |
| C 17 | 59.8 | 7.0 | 1101 | 9 | CNS0021U |
| C 18 | 59.2 | 7.0 | 903 | 9 | CNS0011C |
| C 19 | 59.2 | 7.0 | 1036 | 9 | CNS00599 |
| C 20 | 59 | 6.9 | 1101 | 9 | CNS00E07 |
| C 21 | 58.8 | 6.9 | 1101 | 9 | CNS000D1 |
| C 22 | 58.2 | 6.8 | 1101 | 9 | CNS0001B |
| C 23 | 58 | 6.8 | 1200 | 9 | CNS016CO |
| C 24 | 57.8 | 6.8 | 938 | 9 | CNS00DKY |

| | | | | | |
|------|------|-----|------|---|----------|
| C 25 | 57.8 | 6.8 | 1101 | 9 | CNS00EVL |
| C 26 | 57.8 | 6.8 | 1101 | 9 | CNS0100X |
| C 27 | 57.6 | 6.8 | 1048 | 5 | BX343665 |
| C 28 | 57.4 | 6.7 | 603 | 9 | CNS01115 |
| C 29 | 57.4 | 6.7 | 1442 | 8 | CC303074 |
| C 30 | 57.2 | 6.7 | 879 | 9 | CNS01JRG |
| C 31 | 57 | 6.7 | 973 | 9 | CNS071LE |
| C 32 | 57 | 6.7 | 1352 | 9 | AG382165 |
| C 33 | 56.8 | 6.7 | 746 | 8 | BH921242 |
| C 34 | 56.8 | 6.7 | 1608 | 9 | CL118721 |
| C 35 | 56.6 | 6.7 | 514 | 6 | CA856511 |
| C 36 | 56.6 | 6.7 | 928 | 9 | CNS00DKY |
| C 37 | 56.4 | 6.6 | 1135 | 9 | AG278055 |
| C 38 | 56.4 | 6.6 | 1201 | 9 | CNS015X9 |
| C 39 | 56.4 | 6.6 | 1749 | 9 | AG310735 |
| C 40 | 56.2 | 6.6 | 1067 | 6 | CD384866 |
| C 41 | 56.2 | 6.6 | 1101 | 9 | CNS017KX |
| C 42 | 56 | 6.6 | 984 | 9 | CL113045 |
| C 43 | 56 | 6.6 | 1805 | 9 | CL080711 |
| C 44 | 55.8 | 6.6 | 974 | 9 | CNS001TT |
| C 45 | 55.6 | 6.5 | 1493 | 9 | CL078589 |

ALIGNMENTS

RESULT 1
CNS0039G/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CNS0039G 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR08K10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

AL063921 GI:4941778

GSS.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :

BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Oosagawa and

Aaron Mammossier in Pieter de Jong's laboratory in the Department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain v2; cn bw sp, the same strain used for the BDGP's

pl and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1..1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR08K10"

/clone_lib="RPCI-98"

/note="end : TET3"

ORIGIN

Query Match 9.8%; Score 83.4; DB 9; Length 1101;

Best Local Similarity 17.2%; Pred. No. 7e-08;


```

Db      331 NNNNAATATAAAAAAATATATANAATTTATAA---TATTTTTTATTTTTTATTTT 387
QY      637 ACGGGTCAAGGAAGTGGAGCAAGCAAGCTGGAGTTACTTCTTATAAGAAATGTTTATACT 696
Db      388 ATATTTTAAAAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 447
QY      697 AAATGAGAATGCAACACAGAAATGAGAATGATATATGATTATGTTTAAATATAGTG 756
Db      448 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 507
QY      757 TATTTTATTTTAAAAAGATCGCATACATTACCAGCCAGATGAGTTATTTATCATCACTCA 816
Db      508 TTTTNTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 567
QY      817 CAACAAAGTACAAGA 832
Db      568 ANAANAATAAANAANA 583

CNS0167M      1201 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN15M24 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION     AL106396
VERSION        AL106396.1 GI:5621701
KEYWORDS       GSS
SOURCE         Drosophila melanogaster (fruit fly)
ORGANISM       Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1201)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT        Determination of this BAC-end sequence was carried out as part of a
                collaboration with the European Drosophila Genome Project (EDGP) -
                http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
                library (Dros BAC) was made by Alain Billaud at CSEH (Centre
                d'Etude du Polymorphisme Humain) with funding provided by a MRC
                project grant. The DNA was prepared from embryos by Alain Bucheton
                and Genevieve Payan. It has been constructed in the vector
                pBelobAC11.
FEATURES       source
                1..1201
                /organism="Drosophila melanogaster"
                /mol_type="genomic DNA"
                /db_xref="taxon:7227"
                /clone="BACN15M24"
                /clone_lib="DrosBAC"
                /plasmid="pBelobAC11"
                /note="end : T7"

ORIGIN
Query Match      7.4%; Score 62.8; DB 9; Length 1201;
Best Local Similarity 35.7%; Pred. No. 0.0024;
Matches 200; Conservative 63; Mismatches 297; Indels 0; Gaps 0;

QY      285 AATGAGAAAATTCAGCAACACAGTACAAGAGCTCTGGAAGTAAATGTTGAGAGTATGC 344
Db      1201 ANAWATWATATAWATATATAATAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1142
QY      345 AATAAGGAGATACGGGGATACGCTCAAGTATTTTTCAGCACATAGAAAAGGCAGAA 404
Db      1141 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1082
QY      405 AGGGATAAAGTGCTCTATGAATGGTGAAGTATGGGACCGTTAGTTAGCTTAAAAAAA 464
Db      1081 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1022

```

```

QY      465 GTGACTTCTTACTTGAAGTAATGAAGTGGAGTAGAAGTGAAGTAAAGTAAAGTAAATATA 524
Db      1021 WAWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 962
QY      525 AGTTATTTAAAGTGTGTTGGAAAAAGAAATAGAAGTGTGTAAGAAAAAGTGTAGCATTTTCTACT 584
Db      961 WMTTTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 902
QY      585 TCCAACTTATTTCTCAGCACTTCTTAAAGAGTACTTCTTACTTTTTCACCAACCGGTCA 644
Db      901 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 842
QY      645 AGCAAGTGGAAAGCAAAAAAGCTGGAGTTACTTCTTATAAGAATGTTTATATACTAAATGAGA 704
Db      841 AAAAAATTTTAAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 782
QY      705 AATGACAACACAGAAATGAGAAATGAATATGATTAATGTTTAAATAATAGTGTATTTTAT 764
Db      781 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 722
QY      765 TTAAAAAGATCCATACATTACCAGCCAGATGAAGTTATTCATCACTCACTCAACAAAG 824
Db      721 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 662
QY      825 TAAAAAGAAAAAAGTTGCAAT 844
Db      661 MAMMAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA

CNS00KAE      1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR39P05 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION     AL077628
VERSION        AL077628.1 GI:4956903
KEYWORDS       GSS
SOURCE         Drosophila melanogaster (fruit fly)
ORGANISM       Drosophila melanogaster
                Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
                Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
                Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 1101)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
                BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
                - Web : www.genoscope.cns.fr)
COMMENT        Determination of this BAC-end sequence was carried out as part of a
                collaboration with the Berkeley Drosophila Genome Project (BDGP).
                The BDGP is constructing a physical map of the Drosophila
                melanogaster genome using these BACs. For further information
                please see http://www.fruitfly.org The BDGP Drosophila
                melanogaster BAC library was prepared by Kazuoto Osoegawa and
                Aaron Mammosier in Pieter de Jong's laboratory in the Department of
                Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
                NY. The library is named RPCI-98 and was constructed by partial
                EcoRI digestion of Drosophila DNA provided by the BDGP from the
                isogenic strain y2; cn bw sp, the same strain used for the BDGP's
                p1 and EST libraries. A more detailed description of the library
                and how to order individual BAC clones, the entire library, or
                filters for hybridization from the BACPAC Resource Center can be
                found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES       source
                1..1101
                /organism="Drosophila melanogaster"
                /mol_type="genomic DNA"
                /db_xref="taxon:7227"
                /clone="BACR39P05"
                /clone_lib="RPCI-98"
                /note="end : TET3"

```

```
ORIGIN
Query Match      7.4%; Score 62.6; DB 9; Length 1101;
Best Local Similarity 38.6%; Pred. No. 0.0027;
Matches 245; Conservative 19; Mismatches 365; Indels 5; Gaps 1;

Qy 138 TTGTGATGACTTCTTCATATTTATTGAGAAATACAACTTTGCAGAAAAGCCCACTCTGGG 197
Db 36 TATTATATATATTTANATATNTATANTANANTATATAANTNTNANNTNNTTATATAAAN 95

Qy 198 TTTCAGGTTCTGGAATTAATTAAGCTTTTATCATCACTATGGAAGTGATCTATTTCGAAGC 257
Db 96 TCTNCTATATAGATATATATATATNTNTNANTATATATANTATNTNANANNTTATNA 155

Qy 258 GGAGTGGAGGAACCTTGAGTACGTGAATGAGAAAATTCAGCAACACAGTACAAAGAA 317
Db 156 TATATNANTATANTANANANTANANATATATAANANTNTNTANTNTATAGWGNW 215

Qy 318 GCTCTGGAAGTAATCTTGAGAGTATGCAATAGGAGATACGGGATACGTCAGAGTATT 377
Db 216 NATATANATANTNTNANTNTNTNTNTNTATATATATATANTNANANATATANANNTAA 275

Qy 378 TTTTACGACACAATAGAAAAGCGAAGGGATAAAGTCTCTATGAATGGTGAGGAATT 437
Db 276 ATANTANANATAGWGWATTTATTATATAAAAATAATAATATATATATATATAATT 335

Qy 438 GGGACGGTTTAGTGTAGCTTAAAAAAGTGACTTCTTCTGAAAGTAATGAAGTGGAGTA 497
Db 336 ATAATAAAATTTAATTTTAAAAAATAATTTATTTATATNAAAAATTTATTATATTA 395

Qy 498 GAACCTGATAAGTAAAGTAAATATATTAAGTGTATTAAGTGTTCGAAAAGAAATAGAGT 557
Db 396 AAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 450

Qy 558 TGTAAGAAAAGTTAGCATTTTCTACTTCCAACTTTTCTCAGCACTTCTTAAAAAGTAC 617
Db 451 TTAATAATAATAATATTTTAWAAATTTAATAATATATATATATATATATATATATATAT 510

Qy 618 TTCTTACTTTTACAAACGGGTCAAGGAAAGTGGAGCAAGCAAGCAAGCAAGCAAGCA 677
Db 511 AATANAATAAATAAATTTAAAAATTAAWAAATTAATAAWAAATTAATAAWAAATTAATA 570

Qy 678 TTATAAGATGTTTATCTAATGAGAAATGACAAACACAGAAATGCAATATATATAT 737
Db 571 AWAAWAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 630

Qy 738 TATTGGTTTAATAATAGTGTATTTTATTAAAAA 771
Db 631 AATAAAAAATAAAAAATAAATAAATAAATAAATAA 664

RESULT 9
AG321083/c
LOCUS      1380 bp      DNA      linear      GSS 02-JUN-2004
DEFINITION Mus musculus molossinus DNA, clone:MSMg01-107D08.TJ, genomic survey
sequence.
ACCESSION  AG321083
VERSION     AG321083.1   GI:47894040
KEYWORDS   GSS.
SOURCE     Mus musculus molossinus
ORGANISM   Mus musculus molossinus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1  Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
   BAC end Sequences of Library MSMg01
   JOURNAL   Unpublished
2  (bases 1 to 1380)
   Hattori,M., Toyoda,A., Noguchi,H., Kojima,T. and Sakaki,Y.
   Direct Submission
   Submitted (17-NOV-2003) Masahira Hattori, The Institute of Physical
   and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
   1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

COMMENT
(E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the mouse BAC library MSMg01. For BAC
library availability, please contact Kuniya Abe (abe@rtc.riken.jp).
Tsukuba Institute, Bio Resource Center,
The Institute of Physical and Chemical Research (RIKEN) 3-1-1
Koyadai, Tsukuba, 305-0074 Japan
phone: 81-298-36-9189, fax: 81-298-36-9199
e-mail: abe@rtc.riken.jp
PRIMERS
Sequencing : TJ
LIBRARY      : pBACE3.6
Vector       : EcoRI
R.Site 1    : EcoRI.
R.Site 2    : EcoRI.
FEATURES
    source
        1..1380
            /organism="Mus musculus molossinus"
            /mol_type="genomic DNA"
            /sub_species="molossinus"
            /db_xref="taxon:57486"
            /clone="MSMg01-107D08.TJ"
            /sex="male"
            /tissue_type="mixture of kidney and spleen"
            /clone_lib="MSMg01 Mouse Male BAC Library"
ORIGIN
Query Match      7.3%; Score 61.8; DB 9; Length 1380;
Best Local Similarity 45.4%; Pred. No. 0.0041;
Matches 260; Conservative 0; Mismatches 312; Indels 1; Gaps 1;

Qy 264 GAGGAAGAACTTGAGTACGTGAATGAGAAAAATTCAGCAACACAGTACAAAGAGCTCTG 323
Db 1100 GAGGAAGAGAAAAAGAAAGAGAGAAAAAATAATGTAGAGAAAAATTAAGAGAAAGGAAA 1041

Qy 324 GAAGTAAATGTTGACAGTATGCAATAAGGAGAAATACGGGGATACCTCAAGATATTTTAC 383
Db 1040 AAAGATGAGAGTGTGTGAAGAAAAAGAAAAAGAAAAAGTAAAG-AGAGGAAATATAAA 982

Qy 384 GACCAATAGAAAAAGGAGAAAGGATAAAGTGCTCTATGAATCGTGAGGAAATTTGGGACG 443
Db 981 GTAAAAATAAGAAATAAAGAGAGAGAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAG 922

Qy 444 GTTTAGTGTAGCTTAAAAAAGTGACTTCTTCTTGAAGTAATGCAAGTGGAGTGAAGCTG 503
Db 921 AAGAAAAAAGGTAGAGAGAGAGAAAAAGAAATATGAAAAAAGAAAAAATATATAAGAGAG 862

Qy 504 ATAAGTAAAGTAATAATTATAAGTTATTAAGTCTTTTGGAAAAAGAAATAGAAAGTTGTA 563
Db 861 AGAGAGAAAGTACAGAAAGAGAGAGAAATAATAAGATATAAAGAGAAAGATATAAAG 802

Qy 564 GAAAAGTTAGCAATTTTCTTCCAACTTTATTTCTCAGACTTCTTTAAAAAGTACTTCTTA 623
Db 801 GAAAAGTAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 742

Qy 624 CTTTTTACAAACGGGTCAAGGAAAGTGAAGCAAAAGCTCGAGTACTTCTTATAA 683
Db 741 AAGAAATGTAAAAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 682

Qy 684 GAATGTTTATCTAAATGAGAAATGACAAACACAGAAATGAGAAATGAAATGATATATTGG 743
Db 681 GAAATTAATAGAAAAAGAAATAGTGAAGAAAAAAGAAAAAAGAAATGAAATGAAATGAA 622

Qy 744 TTTAATAATAGTGTATTTTATTTAAAAAGATCGCATATTCACGCCAGATGAAGTTAT 803
Db 621 AATAAGAAAGTAAAAAAGTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 562

Qy 804 TCATCAACTCAACAAAGTACAAAGAAAA 836
Db 561 TCCAAAAACAAAAACAAAAAGCAACAAAA 529

RESULT 10
```


CNS016LI/c
LOCUS
DEFINITION
CNS016LI 1101 bp DNA linear GSS 26-JUL-1999
Drosophila melanogaster genome survey sequence T7 end of BAC
BACN16D22 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBAC11.
FEATURES
source
Location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACN16D22"
/clone_lib="DrosBAC"
/plasmid="pBelOBAC11"
/note="end : T7"
ORIGIN
Query Match 7.2%; Score 61.2; DB 9; Length 1101;
Best Local Similarity 32.8%; Pred. No. 0.0054;
Matches 107; Conservative 76; Mismatches 145; Indels 0; Gaps 0;
QY 509 TAAAGTAATAATATTAAGTTATTAAAGTGTTCGAAAGAAATAGAAAGTTTAAAGAAA 568
DB 1077 TMAVWWVCTMTMAMTAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 1018
QY 569 GTTAGCATTTCTTCACTTATTTCTCAGCACTTCTTAAAGTACTCTTCACTTTT 628
DB 1017 TWTWCAATAHWAHWTATATATWAAHTATTWATAHTATTWATWTTWTTWTTWTTWTT 958
QY 629 TTACACAAACGGGTCAAGGAAAGTGAAGCAAAAGCTGGAGTTACTTCTTATAAGATG 688
DB 957 TWAAWAAWAAWATATAAAYWHAATWAAWAAWAAATTTWTTWTTWTTWTTWTTWTTWTT 898
QY 689 TTTTACTACTAATCAGAAATGACACAAACACAGAAATGAAATGATGATTATTTGGTTTAA 748
DB 897 ATWAAWTAATWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 838
QY 749 TAATAGTGTATTTTATTTAAAGATCGCATACATTACCAGCCAGATGAAGTTATTTCATC 808
DB 837 AWTWAAAAATATATWTTWAAAAAATAWAAWAAWAAATATWAAAAAATAWAAAAAHT 778
QY 809 ACAACTCACAACAAAGTACAAAGAAAA 836
DB 777 AAAAFAAAAAWAAWAAWAAAAAATAA 750
RESULT 11
CNS0039L/c
LOCUS
DEFINITION
CNS0039L 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR08I10 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

AL063926
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .1101
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR08I10"
/clone_lib="RPCI-98"
/note="end : T7"
ORIGIN
Query Match 7.1%; Score 60.8; DB 9; Length 1101;
Best Local Similarity 30.4%; Pred. No. 0.0066;
Matches 127; Conservative 109; Mismatches 177; Indels 5; Gaps 1;
QY 151 TCGATTTTATTCAGAAATTAACAACCTTTGCAGAAAGCAACTCTCGGTTTTCAGGTTCTCG 210
DB 1099 TDKTKTKRWGAGDRDRDRGTRKTRKTRAGARARARARARARARARARARARARARAR 1040
QY 211 AATTAATTAAGCTTTTACATCACTATGGAAGTGATCTATTCCAGCCGAGTGGAGGAAG 270
DB 1039 WAWRAAWNDWGATGDTWTADWDGADDRRAADRRRAADATARKKTKGTGTA----- 985
QY 271 AACTTGAAGTACGTGAATGAGAAAAATTGAGCAACACAGTACAGAAAGCTCTGGAAGTAA 330
DB 984 RATRTKTRTWTWAWAAGTTRATKMGDTDAATWTGTTTARTATGKGTATWKGTGKTT 925
QY 331 TGTTCAGAGTATGCAATAGGAGAATACGGGATACGTCAGAAAGTATTTTTCAGCACAAA 390
DB 924 KKTITKGGTKTKTKAKAWAAAAAATAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAW 865
QY 391 TAGAAAAAGCAGAAAGGATAAAGTCTCTATGAATGGTGAGAAATTTGGACGGTTTAGG 450
DB 864 TKADATADWAWADATKRTKTATRAATTTDKAAAAAATAATGCAAAAAARAAAAAARAKW 805
QY 451 TTAGCTTAAAAAAGTGACTTCTTACTTGAAGTAATGAAGTGGAGTGAAGTGAAGTA 510
DB 804 TDDRWAAARWDAARDARGTGGTAARDGWWGKKAATKWTGRTTAAAAADWAAAAAATAW 745
QY 511 AAGTAATAATTAAGTATTAAAGTGTTTGGAAAAAATAAGTGTGTAAGAAAA 568
DB 744 AWWAAATAATAAATAAATARWGWAAAAATKWTGAAAAAADAWAAARAGDKTTAAAAAAA 687
RESULT 12

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 16:35:35 ; Search time 138.012 Seconds
(without alignments)
10279.455 Million cell updates/sec

Title: US-09-806-197-24
Perfect score: 30
Sequence: 1 aacaatgatgtccgggtacatcatgac 30

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb ba:*
- 2: gb htg:*
- 3: gb in:*
- 4: gb om:*
- 5: gb ov:*
- 6: gb pat:*
- 7: gb ph:*
- 8: gb pl:*
- 9: gb pr:*
- 10: gb ro:*
- 11: gb sts:*
- 12: gb sy:*
- 13: gb un:*
- 14: gb vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------------|--------------------|
| 1 | 30 | 100.0 | 2048 | 6 E40089 | E40089 Plant promo |
| 2 | 30 | 100.0 | 2048 | 6 E40090 | E40090 Plant promo |
| 3 | 30 | 100.0 | 2052 | 6 E40087 | E40087 Plant promo |
| 4 | 30 | 100.0 | 2052 | 6 E40093 | E40093 Plant promo |
| 5 | 30 | 100.0 | 2056 | 6 E40091 | E40091 Plant promo |
| 6 | 23.6 | 78.7 | 2042 | 6 AR076817 | AR076817 Sequence |
| 7 | 23.6 | 78.7 | 2042 | 6 E15125 | E15125 Promoter. 7 |
| 8 | 21.8 | 72.7 | 348283 | 2 AC125706 | AC125706 Rattus no |
| 9 | 21.2 | 70.7 | 161762 | 2 CR352244 | CR352244 Danio rer |
| 10 | 21.2 | 70.7 | 213033 | 2 AC110515 | AC110515 Mus muscu |
| 11 | 21 | 70.0 | 140142 | 8 AC146937 | AC146937 Oryza sat |
| 12 | 21 | 70.0 | 185386 | 10 AC025586 | AC025586 Genomic s |
| 13 | 21 | 70.0 | 205668 | 10 AL592422 | AL592422 Mouse DNA |
| 14 | 20.6 | 68.7 | 69674 | 9 AL136130 | AL136130 Human DNA |
| 15 | 20.6 | 68.7 | 78026 | 2 AC023516 | AC023516 Homo sapi |
| 16 | 20.6 | 68.7 | 215896 | 2 AC046180 | AC046180 Homo sapi |
| 17 | 20.4 | 68.0 | 118001 | 9 AC073081 | AC073081 Homo sapi |
| 18 | 20.4 | 68.0 | 193910 | 2 AC040996 | AC040996 Homo sapi |
| 19 | 20.4 | 68.0 | 194235 | 9 AC109822 | AC109822 Homo sapi |

| | | | | | | | |
|---|----|------|------|--------|----|-----------|--------------------|
| c | 20 | 20.4 | 68.0 | 288385 | 2 | AC092430 | Homo sapi |
| c | 21 | 20.2 | 67.3 | 121101 | 2 | AC008034 | Homo sapi |
| c | 22 | 20.2 | 67.3 | 155313 | 2 | AC068315 | Homo sapi |
| c | 23 | 20.2 | 67.3 | 157007 | 2 | AC066599 | Homo sapi |
| c | 24 | 20.2 | 67.3 | 161903 | 2 | AC021996 | Homo sapi |
| c | 25 | 20.2 | 67.3 | 185067 | 9 | AC022382 | Homo sapi |
| c | 26 | 20.2 | 67.3 | 189430 | 2 | AC011610 | Homo sapi |
| c | 27 | 20.2 | 67.3 | 197360 | 2 | AC018829 | Homo sapi |
| c | 28 | 20.2 | 67.3 | 216530 | 2 | BX649535 | Danio rer |
| c | 29 | 20.2 | 67.3 | 260792 | 2 | AC125770 | Rattus no |
| c | 30 | 20.2 | 67.3 | 320398 | 2 | AC136192 | Rattus no |
| c | 31 | 20 | 66.7 | 81398 | 2 | AC023821 | Homo sapi |
| c | 32 | 20 | 66.7 | 103152 | 2 | AP004091 | Oryza sat |
| c | 33 | 20 | 66.7 | 142010 | 2 | AC139637 | Mus muscu |
| c | 34 | 20 | 66.7 | 159601 | 8 | OSJN00059 | AL606622 Oryza sat |
| c | 35 | 20 | 66.7 | 163328 | 8 | AP004865 | Oryza sat |
| c | 36 | 20 | 66.7 | 172879 | 10 | AC142474 | Mus muscu |
| c | 37 | 20 | 66.7 | 201294 | 2 | AC128748 | Rattus no |
| c | 38 | 20 | 66.7 | 204302 | 2 | AC068650 | Mus muscu |
| c | 39 | 20 | 66.7 | 216935 | 2 | AC132648 | Rattus no |
| c | 40 | 20 | 66.7 | 234787 | 2 | AC073780 | Mus muscu |
| c | 41 | 20 | 66.7 | 244520 | 2 | AC128555 | Rattus no |
| c | 42 | 20 | 66.7 | 245754 | 2 | AC127884 | Rattus no |
| c | 43 | 20 | 66.7 | 254067 | 2 | AC133699 | Rattus no |
| c | 44 | 20 | 66.7 | 266396 | 2 | AC108524 | Rattus no |
| c | 45 | 20 | 66.7 | 315370 | 2 | AC109390 | Rattus no |

ALIGNMENTS

| | | | | | |
|-----------------------|---|--------------------------------|-------|--------------|-----------------|
| RESULT 1 | E40089 | 2048 bp | DNA | linear | PAT 31-JAN-2002 |
| LOCUS | E40089 | Plant promoter and terminator. | | | |
| DEFINITION | E40089 | | | | |
| ACCESSION | E40089.1 | GI:18627205 | | | |
| VERSION | JP 2000166577-A/3. | | | | |
| KEYWORDS | unidentified | | | | |
| SOURCE | unidentified | | | | |
| ORGANISM | unclassified. | | | | |
| REFERENCE | 1 (bases 1 to 2048) | | | | |
| AUTHORS | Nishikawa,S. and Oeda,K. | | | | |
| TITLE | Plant promoter and terminator | | | | |
| JOURNAL | Patent: JP 2000166577-A 3 20-JUN-2000; | | | | |
| COMMENT | SUMITOMO CHEM CO LTD | | | | |
| | OS Daucus carota L. | | | | |
| | PN JP 2000166577-A/3 | | | | |
| | PD 20-JUN-2000 | | | | |
| | PF 01-OCT-1999 JP 1999281475 | | | | |
| | PR | | | | |
| | PI SATOMI NISHIKAWA,KENJI OEDA | | | | |
| | PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1:91), PC | | | | |
| | C12N15/00, | | | | |
| | PC C12N5/00,(C12N5/00,C12R1:91) | | | | |
| | CC | | | | |
| | FH Key Location/Qualifiers | | | | |
| | FT promoter Location/Qualifiers | | | | |
| | 1..2048 | | | | |
| | /organism="unidentified" | | | | |
| | /mol_type="genomic DNA" | | | | |
| | /db_xref="taxon:32644" | | | | |
| FEATURES | source | | | | |
| ORIGIN | | | | | |
| Query Match | 100.0%; | Score 30; | DB 6; | Length 2048; | |
| Best Local Similarity | 100.0%; | Pred. No. 0.00047; | | | |
| Matches | 30; | Conservative | 0; | Mismatches | 0; |
| | | | | Indels | 0; |
| | | | | Gaps | 0; |
| QY | 1 AACATGATGTCCGGGTACATCATGAC 30 | | | | |
| DB | 1742 AACATGATGTCCGGGTACATCATGAC 1771 | | | | |

```
RESULT 2
LOCUS           E40090           2048 bp          DNA          linear          PAT 31-JAN-2002
DEFINITION      Plant promoter and terminator.
ACCESSION       E40090
VERSION         E40090.1 GI:18627206
KEYWORDS        JP 2000166577-A/4.
SOURCE          unidentified
ORGANISM        unclassified.
REFERENCE       1 (bases 1 to 2048)
AUTHORS        Nishikawa, S. and Oeda, K.
TITLE          Plant promoter and terminator
JOURNAL        Patent: JP 2000166577-A 4 20-JUN-2000;
               SUMITOMO CHEM CO LTD
COMMENT        OS Daucus carota L.
               PN JP 2000166577-A/4
               PD 20-JUN-2000
               PF 01-OCT-1999 JP 1999281475
               PR
               PI SATOMI NISHIKAWA, KENJI OEDA
               PC C12N15/09, A01H5/00, C12N1/21, C12N5/10//((C12N5/10, C12R1:91), PC
                  C12N15/00,
               CC C12N5/00, (C12N5/00, C12R1:91)
               FH Key Location/Qualifiers
               FT promoter (1)..(2048).
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"

FEATURES
   source
ORIGIN
Query Match      100.0%; Score 30; DB 6; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
    |||
Db 1742 AACAAATGATGTCGGGTGTACATCTATGAC 1771

RESULT 3
LOCUS           E40087           2052 bp          DNA          linear          PAT 31-JAN-2002
DEFINITION      Plant promoter and terminator.
ACCESSION       E40087
VERSION         E40087.1 GI:18627203
KEYWORDS        JP 2000166577-A/1.
SOURCE          unidentified
ORGANISM        unclassified.
REFERENCE       1 (bases 1 to 2052)
AUTHORS        Nishikawa, S. and Oeda, K.
TITLE          Plant promoter and terminator
JOURNAL        Patent: JP 2000166577-A 1 20-JUN-2000;
               SUMITOMO CHEM CO LTD
COMMENT        OS Daucus carota L.
               PN JP 2000166577-A/1
               PD 20-JUN-2000
               PF 01-OCT-1999 JP 1999281475
               PR
               PI SATOMI NISHIKAWA, KENJI OEDA
               PC C12N15/09, A01H5/00, C12N1/21, C12N5/10//((C12N5/10, C12R1:91), PC
                  C12N15/00,
               CC C12N5/00, (C12N5/00, C12R1:91)
               FH Key Location/Qualifiers
               FT promoter (1)..(2052).
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"

FEATURES
   source
```

```
ORIGIN
Query Match      100.0%; Score 30; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
    |||
Db 1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775

RESULT 4
LOCUS           E40093           2052 bp          DNA          linear          PAT 31-JAN-2002
DEFINITION      Plant promoter and terminator.
ACCESSION       E40093
VERSION         E40093.1 GI:18627209
KEYWORDS        JP 2000166577-A/7.
SOURCE          unidentified
ORGANISM        unclassified.
REFERENCE       1 (bases 1 to 2052)
AUTHORS        Nishikawa, S. and Oeda, K.
TITLE          Plant promoter and terminator
JOURNAL        Patent: JP 2000166577-A 7 20-JUN-2000;
               SUMITOMO CHEM CO LTD
COMMENT        OS Daucus carota L.
               PN JP 2000166577-A/7
               PD 20-JUN-2000
               PF 01-OCT-1999 JP 1999281475
               PR
               PI SATOMI NISHIKAWA, KENJI OEDA
               PC C12N15/09, A01H5/00, C12N1/21, C12N5/10//((C12N5/10, C12R1:91), PC
                  C12N15/00,
               CC C12N5/00, (C12N5/00, C12R1:91)
               FH Key Location/Qualifiers
               FT promoter (1)..(2052).
               /organism="unidentified"
               /mol_type="genomic DNA"
               /db_xref="taxon:32644"

FEATURES
   source
ORIGIN
Query Match      100.0%; Score 30; DB 6; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTGTACATCTATGAC 30
    |||
Db 1746 AACAAATGATGTCGGGTGTACATCTATGAC 1775

RESULT 5
LOCUS           E40091           2056 bp          DNA          linear          PAT 31-JAN-2002
DEFINITION      Plant promoter and terminator.
ACCESSION       E40091
VERSION         E40091.1 GI:18627207
KEYWORDS        JP 2000166577-A/5.
SOURCE          unidentified
ORGANISM        unclassified.
REFERENCE       1 (bases 1 to 2056)
AUTHORS        Nishikawa, S. and Oeda, K.
TITLE          Plant promoter and terminator
JOURNAL        Patent: JP 2000166577-A 5 20-JUN-2000;
               SUMITOMO CHEM CO LTD
COMMENT        OS Daucus carota L.
```

```

PN JP 2000166577-A/5
PD 20-JUN-2000
PF 01-OCT-1999 JP 1999281475
PI SATOMI NISHIKAWA, KENJI OEDA
PC C12N15/09,A01H5/00,C12N1/21,C12N5/10/(C12N5/10,C12R1:91), PC
C12N15/00,
PC C12N5/00,(C12N5/00,C12R1:91)
CC Key Location/Qualifiers
FH promoter (1).(2056).
FT Location/Qualifiers
FEATURES
source
1..2056
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 100.0%; Score 30; DB 6; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0.00047;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACATGTATGTCGGGTGACATCTATGAC 30
Db 1746 AACATGTATGTCGGGTGACATCTATGAC 1775
|||||
RESULT 6
AR076817
LOCUS AR076817 2042 bp DNA linear PAT 30-AUG-2000
DEFINITION Sequence 2 from patent US 5959176.
ACCESSION AR076817
VERSION AR076817.1 GI:10003563
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,S. and Oeda,K.
TITLE Plant promoter and utilization thereof
JOURNAL Patent: US 5959176-A 2 28-SEP-1999;
FEATURES
source
1..2042
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 78.7%; Score 23.6; DB 6; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.92;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AACATGTATGTCGGGTGACATCTATGAC 30
Db 1737 AACACGTTTGTCGGGTGATATTATGAC 1766
|||||
RESULT 7
E15125
LOCUS E15125 2042 bp DNA linear PAT 28-JUL-1999
DEFINITION Promoter.
ACCESSION E15125
VERSION E15125.1 GI:5709808
KEYWORDS JP 1998052273-A/2.
SOURCE unidentified
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2042)
AUTHORS Torikai,T. and Oita,K.
TITLE VEGETABLE PROMOTER AND ITS USE
JOURNAL Patent: JP 1998052273-A 2 24-FEB-1998;
SUMITOMO CHEM CO LTD
COMMENT OS Daucus carota L. (carrot)
PN JP 1998052273-A/2

```

```

PD 24-FEB-1998
PF 12-AUG-1996 JP 1996212680
PI TORIKAI TOSHIMI, OITA KENJI
PC C12N15/09,A01H5/00,C07H21/04,C07K14/415,C12N1/21,C12N5/10; CC
strandedness: Single;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2042
/organism="Daucus carota L."
FT promoter 1<..<2042.
FT Location/Qualifiers
FEATURES
source
1..2042
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
ORIGIN
Query Match 78.7%; Score 23.6; DB 6; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.92;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 AACATGTATGTCGGGTGACATCTATGAC 30
Db 1737 AACACGTTTGTCGGGTGATATTATGAC 1766
|||||
RESULT 8
AC125706
LOCUS AC125706 348283 bp DNA linear HTG 13-NOV-2002
DEFINITION Rattus norvegicus clone CH230-1109, *** SEQUENCING IN PROGRESS ***,
8 unordered pieces.
ACCESSION AC125706
VERSION AC125706.5 GI:24941053
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 348283)
AUTHORS Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allien,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,B., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hognes,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jollivet,A.,
Karpathy,S., Kelly,S., Khan,Z., King,L., Kovar,C., King,L., Liu,J.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozano,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhinney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munitasa,M., Murphy,M., Nair,L.,

```

Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,
 Nwokedemeh, O., Okwundu, G., Olarinmoye, A., Pal, S., Parks, K.,
 Paternek, S., Paul, H., Perez, A., Perez, L., Pfamkoch, C.,
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Shen, H.,
 Sanders, M., Savery, G., Scherer, S., Scott, G., Shateman, S., Shen, H.,
 Shetty, J., Sivatsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D.,
 Sneed, A., Sodergren, E., Sung, X.-Z., Sotelle, R., Sosa, J.,
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J.,
 Wang, Q., Wang, S., Warren, J., Warren, K., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

11765: contig of 11765 bp in length
 11865: gap of unknown length
 48867: contig of 37002 bp in length
 48868: gap of unknown length
 54443: contig of 5476 bp in length
 54543: gap of unknown length
 59602: contig of 5059 bp in length
 59702: gap of unknown length
 343138: contig of 283436 bp in length
 343238: gap of unknown length
 344661: contig of 1423 bp in length
 344761: gap of unknown length
 345971: contig of 1210 bp in length
 346072: gap of unknown length
 346283: contig of 2212 bp in length.
 Location/Qualifiers
 source
 1. 348283
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-11J9"
 misc_feature
 1. 1925
 /note="wgs end extension
 clone_end:Sp6"
 misc_feature
 1909..2527
 /note="clone boundary
 clone_end:Sp6
 site:
 end sequence:BH342645"
 8272..11765
 /note="wgs contig"
 11866..15419
 /note="wgs contig"
 20601..22623
 /note="wgs contig"
 22674..27148
 /note="wgs contig"
 48968..52962
 /note="wgs contig"
 53013..54443
 /note="wgs contig"
 54544..56816
 /note="wgs contig"
 56870..59602
 /note="wgs contig"
 59703..61990
 /note="wgs contig"
 68117..71973
 /note="wgs contig"
 338673..340201
 /note="wgs contig"
 340622..343138
 /note="wgs contig"
 ORIGIN

Query Match 72.7%; Score 21.8; DB 2; Length 348283;
 Best Local Similarity 92.0%; Pred.No.12;
 Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGGGTACATCT 25
 Db 90808 AACAAATGATGTCGGGTACATCT 90832

RESULT 9

CR352244/c

LOCUS

DEFINITION

CR352244

unordered pieces.

ACCESSION

CR352244

VERSION

HTG; HTGS PHASE1.

KEYWORDS

Danio rerio (zebrafish)

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 8 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)
 * NOTE: This sequence may represent more than one clone.
 * NOTE: This is a 'working draft' sequence. It currently
 consists of 8 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

ORGANISM

Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 161762)

REFERENCE
AUTHORS

Simb,S.

TITLE
JOURNAL

Direct Submission
Submitted (26-MAR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Mar 27, 2004 this sequence version replaced gi:45598726.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: zC131E11
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 157149 bases at least Q40
Consensus quality: 157757 bases at least Q30
Consensus quality: 158375 bases at least Q20
Insert size: 160462; sum-of-contigs
Insert size: 174990; 4.8% error; agarose-fp
Quality coverage: 8.14x in Q20 bases; sum-of-contigs Quality
coverage: 7.67x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 16722: contig of 16722 bp in length
* 16723 16822: gap of 100 bp
* 16823 28393: contig of 11571 bp in length
* 28394 28493: gap of 100 bp
* 28494 40368: contig of 11875 bp in length
* 40369 40468: gap of 100 bp
* 40469 57841: contig of 17373 bp in length
* 57842 57941: gap of 100 bp
* 57942 69071: contig of 11130 bp in length
* 69072 69171: gap of 100 bp
* 69172 77218: contig of 8047 bp in length
* 77219 77318: gap of 100 bp
* 77319 85619: contig of 8301 bp in length
* 85620 85719: gap of 100 bp
* 85720 89373: contig of 3634 bp in length
* 89374 89473: gap of 100 bp
* 89474 98585: contig of 9112 bp in length
* 98586 98685: gap of 100 bp
* 98686 103814: contig of 5129 bp in length
* 103815 103914: gap of 100 bp
* 103915 110517: contig of 6603 bp in length
* 110518 110617: gap of 100 bp
* 110618 139606: contig of 28989 bp in length
* 139607 139706: gap of 100 bp
* 139707 144556: contig of 4850 bp in length
* 144557 144657: gap of 100 bp
* 144657 161762: contig of 17106 bp in length.

FEATURES

source

Location/Qualifiers
1..161762
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-131E11"
/clone_lib="CHORI-211"
1..16722
/note="assembly_fragment:01696"

misc_feature

fragment_chain:1
clone_end:SP6
vector_side:left
16823..28393
/note="assembly_fragment:00818
fragment_chain:1"
28494..40368
/note="assembly_fragment:00996
fragment_chain:1"
40469..57841
/note="assembly_fragment:01195
fragment_chain:1"
57942..69071
/note="assembly_fragment:00668
fragment_chain:1"
69172..77218
/note="assembly_fragment:00429
fragment_chain:1"
77319..85619
/note="assembly_fragment:00216
fragment_chain:1"
85720..89373
/note="assembly_fragment:00050
fragment_chain:1"
89474..98585
/note="assembly_fragment:00542
fragment_chain:1"
98686..103814
/note="assembly_fragment:00090
fragment_chain:1"
103915..110517
/note="assembly_fragment:00319
fragment_chain:1"
110618..139606
/note="assembly_fragment:02011
fragment_chain:2"
139707..144556
/note="assembly_fragment:00147
fragment_chain:2"
144557..161762
/note="assembly_fragment:01414
fragment_chain:2"
clone_end:T7
vector_side:right

ORIGIN

Query Match 70.7%; Score 21.2; DB 2; Length 161762;
Best Local Similarity 88.5%; Pred. NO. 23;
Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 AATGTATGTCGGTGTCATCTATGA 29

Db 9417 AATGTGTCGGTGTCATCTATGA 9392

RESULT 10

AC110515

LOCUS

DEFINITION

AC110515

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC110515 213033 bp DNA linear HTG 09-MAR-2004
Mus musculus chromosome 13 clone RP23-295C1 map 13, WORKING DRAFT
SEQUENCE, 8 ordered pieces.

AC110515.4 GI:44886734
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse)

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 213033)

Birren,B., Nusbaum,C. and Lander,E.
Unpublished
Mus musculus chromosome 13, clone RP23-295C1
2 (bases 1 to 213033)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Canarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kelle, C., LaRoque, K., Lamarca, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 213033)

REFERENCE
AUTHORS
Birken, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K., Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., MacLean, C., MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (09-MAR-2004) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 2, 2004 this sequence version replaced gi:28460874.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WTR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L20363

Center clone name: 295_C1

----- Summary Statistics

Chemistry: Vector: Plasmid; n/a; 100% of reads

Assembly: Program: Phrap; version 0.960731

Consensus quality: 210636 bases at least Q40

Consensus quality: 211310 bases at least Q30

Consensus quality: 211514 bases at least Q20

Insert size: 186000; agarose-fp

Quality coverage: 11.5 in Q20 bases; agarose-fp

Quality coverage: 10.1 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1
* 10790: contig of 10790 bp in length
* 10890: gap of 100 bp
* 66637: contig of 55747 bp in length
* 66737: gap of 100 bp
* 66738: contig of 3141 bp in length
* 69878: gap of 100 bp
* 69978: gap of 100 bp
* 77514: contig of 7536 bp in length
* 77614: gap of 100 bp
* 77615: contig of 11116 bp in length
* 88830: gap of 100 bp
* 88831: contig of 4100 bp in length
* 92930: gap of 100 bp
* 92931: contig of 95814 bp in length
* 188844: gap of 100 bp
* 188845: contig of 24089 bp in length.
* 188945: 213033: contig of 24089 bp in length.

FEATURES
source

1..213033
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="13"
/map="13"
/clone_lib="RPCI-23 Female Mouse BAC"
1..10790
/notes="assembly_fragment
vector_side:left"
10891..66637
/notes="assembly_fragment"
66738..69878
/notes="assembly_fragment"
69879..77514
/notes="assembly_fragment"
77615..88730
/notes="assembly_fragment"
88831..92930
/notes="assembly_fragment"
93031..188844
/notes="assembly_fragment"
188945..213033
/notes="assembly_fragment
clone_end:17"
vector_side:right"

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 70.7%; Score 21.2; DB 2; Length 213033;

Best Local Similarity 88.5%; Pred. No. 23;

Matches 23; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CAATGTATGTCGGGTACATCTATG 28

Db 9398 CTATGTATGTCGGATGTACGTCTATG 9423

RESULT 11

AC146937/c

LOCUS

DEFINITION Oryza sativa chromosome 11 BAC clone OSJNbA0004015, complete

ACCESSION AC146937

VERSION AC146937.2 GI:48958697

KEYWORDS HTG.

SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)

linear

PLN 19-JUN-2004

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS
1 (bases 1 to 140142)
Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansberger, K., Jones, K.M., Overton II, L., Tsitrin, T., Kim, M., Bera, J., Jin, S., Fadrosch, D.W., Tallon, L., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S., Riedmuller, S.B., Utterbach, T., Feidblyum, T., Yang, Q., Haas, B., Suh, B., Peterson, J., Quackenbush, J., White, O., Salzberg, S. and Fraser, C.

TITLE
Oryza sativa ssp. japonica cv. Nipponbare OSJNBA0004015 BAC genomic sequence

JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 140142)
Buell, R.

TITLE
Direct Submission
JOURNAL
Submitted (25-OCT-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

REFERENCE
AUTHORS
3 (bases 1 to 140142)
Buell, R.

TITLE
Direct Submission
JOURNAL
Submitted (30-OCT-2003) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

REFERENCE
AUTHORS
4 (bases 1 to 140142)
Buell, R.

TITLE
Direct Submission
JOURNAL
Submitted (19-JUN-2004) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA

COMMENT
On Jun 19, 2004 this sequence version replaced gi:37991935.

FEATURES
Location/Qualifiers
1..140142
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="11"
/clone="OSJNBA0004015"
/note="japonica cultivar-group"

ORIGIN

Query Match 70.0%; Score 21; DB 8; Length 140142;
Best Local Similarity 82.8%; Pred. No. 28;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 ACAATGTATGTCGGTGTACATCTATGAC 30

Db 113334 ACAATATTTGTTGGTGTACATGTATGAC 113306

RESULT 12
AC025586
LOCUS
DEFINITION
Genomic sequence for Mus musculus, clone RP23-320C8, complete sequence.

ACCESSION
AC025586
VERSION
AC025586.4 GI:21954008
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS
1 (bases 1 to 195386)
McCombie, W.R., de la Bastide, M., Spiegel, L., Preston, R., Nascimento, L., Zutavern, T., Balija, V., Bell, M., Miller, B., Katzenberger, F., Muller, S., Sullivan, P., Yang, C., Dike, S., Palmer, L., O'Shaughnessy, A. and Dedhia, N.

TITLE
Genomic sequence for Mus musculus, clone RP23-320C8, complete sequence

JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 195386)
McCombie, W.R.
Direct Submission
Submitted (13-MAR-2000) Lita Annenberg Hazen Genome Sequencing

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

On Jul 24, 2002 this sequence version replaced gi:99723113
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

FEATURES
source

1..185386
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-320C8"
/clone_lib="RPCI-23"
complement(23305..23539)
/note="We believe the assembly to be correct. The sequence is a simple repeat (TTCCC)n in which the exact number of repeat copies is unknown. One high quality subclone in the region spans the repeat into unique sequence on both sides."

misc_feature

ORIGIN

Query Match 70.0%; Score 21; DB 10; Length 185386;
Best Local Similarity 82.8%; Pred. No. 29;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACAAATGTATGTCGGTGTACATCTATGAC 29

Db 98187 AACAAATGTAAGTCTGGTTAGATCTAAGA 98215

RESULT 13
AL592422
LOCUS
DEFINITION
Mouse DNA sequence from clone RP23-218016 on chromosome 11, complete sequence.

ACCESSION
AL592422
VERSION
AL592422.11 GI:18855205
KEYWORDS
HTG.
SOURCE
Mus musculus (house mouse)

ORGANISM
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Almeida, J.

Direct Submission
Submitted (16-FEB-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Feb 21, 2002 this sequence version replaced gi:17426517.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given

Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
3 (bases 1 to 185386)
McCombie, W.R.

Direct Submission
Submitted (24-JUL-2002) Lita Annenberg Hazen Genome Sequencing Center, Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA

in the feature table with their source databases: Em., EMBL, Sw., SWISSPROT, Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-218016 is from the RPCI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6
This sequence is the entire insert of clone RP23-218016.

FEATURES

```

source
  1..205668
    /location=Qualifiers
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /chromosome="11"
    /clone="RP23-218016"
    /clone_lib="RPCI-23"
    /note="Single clone region. Assembly confirmed by restriction digest data."
misc_feature
  108844
    /note="Random repeat. Forced join. Gap size estimated to be approximately 100bp by restriction digest data."
misc_feature
  108845..108978
    /note="Sequence from uni-directional dGTP big dye terminator reads only"
misc_feature
  121988
    /note="Random repeat. Forced join. Gap size estimated to be approximately 200bp by restriction digest data."

```

ORIGIN

```

Query Match      70.0%; Score 21; DB 10; Length 205668;
Best Local Similarity 82.8%; Pred. No. 29;
Matches 24; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AACATGATATGTCGGGTGATCATCTATGA 29
   ||||| ||||| ||||| ||||| ||||| |||||
Db 183167 AACATGTAAGTCTGTTAGTATCTAAGA 183195

```

RESULT 14

```

AL136130
LOCUS      Human DNA sequence from clone RP1-39M18 on chromosome 6q26-27,
complete sequence.
ACCESSION  AL136130
VERSION    AL136130.7 GI:6911361
KEYWORDS   HTG.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 69674)
  Laird,G.
  Direct Submission
  Submitted (04-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
  Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
  humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
  On Feb 7, 2000 this sequence version replaced gi:678223.
  During sequence assembly data is compared from overlapping clones.
  Where differences are found these are annotated as variations
  together with a note of the overlapping clone name. Note that the
  variation annotation may not be found in the sequence submission
  corresponding to the overlapping clone, as we submit sequences with
  only a small overlap as described above.
  The following abbreviations are used to associate primary accession
  numbers given in the feature table with their source databases:
  Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information
  on the WORMPEP database can be found at
  http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence
  was generated from part of bacterial clone contigs of human
  chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
  Group. Further information can be found at

```

COMMENT

<http://www.sanger.ac.uk/HGP/Chr6>
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
RP1-39M18 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: PCYPAC2

FEATURES

```

source
  1..69674
    /location=Qualifiers
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="6"
    /map="q26-27"
    /clone="RP1-39M18"
    /clone_lib="RPCI-1"

```

ORIGIN

```

Query Match      68.7%; Score 20.6; DB 9; Length 69674;
Best Local Similarity 85.2%; Pred. No. 43;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 ACAATGATATGTCGGGTGATCATCTATG 28
   ||||| ||||| ||||| ||||| ||||| |||||
Db 8130 AAAACGATATGTCAGTGTACATTTATG 8156

```

RESULT 15

```

AC023516
LOCUS      Homo sapiens clone RP11-11504, LOW-PASS SEQUENCE SAMPLING.
ACCESSION  AC023516
VERSION    AC023516.2 GI:9112723
KEYWORDS   HTG; HTGS PHASE0.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 78026)
  Birren,B., Linton,L., Nussbaum,C. and Lander,E.
  Homo sapiens, clone RP11-11504
  Unpublished
REFERENCE  2 (bases 1 to 78026)
  Birren,B., Linton,L., Nussbaum,C., Lander,E., Abraham,H., Allen,N.,
  Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,
  Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
  Choepel,Y., Collangelo,M., Collins,S., Collimore,A., Cooke,P.,
  Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
  Feneator,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
  Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
  Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
  Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
  Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
  Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
  McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
  Meneus,L., Miheva,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
  Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.N.,
  Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
  Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
  Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
  Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Titrell,A.,

```

Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
 Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and
 Zody, M.
 Direct Submission
 Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6978142.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

 Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

 Project Information
 Center project name: L1909
 Center clone name: L15_O_4

* NOTE: This record contains 84 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 800: contig of 800 bp in length
 * 801 900: gap of 100 bp
 * 901 1749: contig of 849 bp in length
 * 1750 1849: gap of 100 bp
 * 1850 2673: contig of 824 bp in length
 * 2674 2773: gap of 100 bp
 * 2774 3614: contig of 841 bp in length
 * 3615 3714: gap of 100 bp
 * 3715 4522: contig of 808 bp in length
 * 4523 4622: gap of 100 bp
 * 4623 5455: contig of 833 bp in length
 * 5456 5555: gap of 100 bp
 * 5556 6388: contig of 833 bp in length
 * 6389 6488: gap of 100 bp
 * 6489 7307: contig of 819 bp in length
 * 7308 7407: gap of 100 bp
 * 7408 8226: contig of 819 bp in length
 * 8227 8326: gap of 100 bp
 * 8327 9160: contig of 834 bp in length
 * 9161 9260: gap of 100 bp
 * 9261 10064: contig of 804 bp in length
 * 10065 10164: gap of 100 bp
 * 10165 10396: contig of 832 bp in length
 * 10397 11099: gap of 100 bp
 * 11097 11910: contig of 814 bp in length
 * 11911 12010: gap of 100 bp
 * 12011 12855: contig of 845 bp in length
 * 12856 12955: gap of 100 bp
 * 12956 13800: contig of 845 bp in length
 * 13801 13900: gap of 100 bp
 * 13901 14730: contig of 830 bp in length
 * 14731 14830: gap of 100 bp
 * 14831 15675: contig of 845 bp in length
 * 15676 15775: gap of 100 bp
 * 15776 16610: contig of 835 bp in length
 * 16611 16710: gap of 100 bp
 * 16711 17539: contig of 829 bp in length
 * 17540 17639: gap of 100 bp
 * 17640 18474: contig of 835 bp in length
 * 18475 18574: gap of 100 bp
 * 18575 19389: contig of 815 bp in length
 * 19390 19489: gap of 100 bp
 * 19490 20328: contig of 839 bp in length

* 20329 20428: gap of 100 bp
 * 20429 21276: contig of 848 bp in length
 * 21277 21376: gap of 100 bp
 * 21377 22174: contig of 798 bp in length
 * 22175 22274: gap of 100 bp
 * 22275 22713: contig of 839 bp in length
 * 22714 23213: gap of 100 bp
 * 23214 24057: contig of 844 bp in length
 * 24058 24157: gap of 100 bp
 * 24158 25006: contig of 849 bp in length
 * 25007 25106: gap of 100 bp
 * 25107 25935: contig of 829 bp in length
 * 25936 26035: gap of 100 bp
 * 26036 26870: contig of 835 bp in length
 * 26871 26970: gap of 100 bp
 * 26971 27812: contig of 842 bp in length
 * 27813 27912: gap of 100 bp
 * 27913 28735: contig of 823 bp in length
 * 28736 28835: gap of 100 bp
 * 28836 29673: contig of 838 bp in length
 * 29674 29773: gap of 100 bp
 * 29774 30596: contig of 823 bp in length
 * 30597 30696: gap of 100 bp
 * 30697 31541: contig of 845 bp in length
 * 31542 31641: gap of 100 bp
 * 31642 32489: contig of 848 bp in length
 * 32490 32589: gap of 100 bp
 * 32590 33378: contig of 789 bp in length
 * 33379 33478: gap of 100 bp
 * 33479 34332: contig of 854 bp in length
 * 34333 34432: gap of 100 bp
 * 34433 35294: contig of 862 bp in length
 * 35295 35394: gap of 100 bp
 * 35395 36239: contig of 845 bp in length
 * 36240 36339: gap of 100 bp
 * 36340 37161: contig of 822 bp in length
 * 37162 37262: contig of 826 bp in length
 * 37262 38087: contig of 826 bp in length
 * 38088 38187: gap of 100 bp
 * 38188 39010: contig of 823 bp in length
 * 39011 39110: gap of 100 bp
 * 39111 39947: contig of 837 bp in length
 * 39948 40047: gap of 100 bp
 * 40048 40886: contig of 839 bp in length
 * 40887 40986: gap of 100 bp
 * 40987 41829: contig of 843 bp in length
 * 41830 41929: gap of 100 bp
 * 41930 42749: contig of 820 bp in length
 * 42750 42849: gap of 100 bp
 * 42850 43687: contig of 838 bp in length
 * 43688 43787: gap of 100 bp
 * 43788 44620: contig of 833 bp in length
 * 44621 44720: gap of 100 bp
 * 44721 45568: contig of 848 bp in length
 * 45569 45669: gap of 100 bp
 * 45669 46495: contig of 827 bp in length
 * 46496 46595: gap of 100 bp
 * 46596 47421: contig of 826 bp in length
 * 47422 47521: gap of 100 bp
 * 47522 48345: contig of 824 bp in length
 * 48346 48445: gap of 100 bp
 * 48446 49286: contig of 841 bp in length
 * 49287 49386: gap of 100 bp
 * 49387 50198: contig of 812 bp in length
 * 50199 50298: gap of 100 bp
 * 50299 51145: contig of 847 bp in length
 * 51146 51245: gap of 100 bp
 * 51246 52064: contig of 819 bp in length
 * 52065 52164: gap of 100 bp
 * 52165 53006: contig of 842 bp in length
 * 53007 53106: gap of 100 bp
 * 53107 53929: contig of 823 bp in length
 * 53930 54029: gap of 100 bp

```

* 54030 54835: contig of 806 bp in length
* 54836 54935: gap of 100 bp
* 54936 55759: contig of 824 bp in length
* 55760 55859: gap of 100 bp
* 55860 56681: contig of 822 bp in length
* 56682 56781: gap of 100 bp
* 56782 57607: contig of 826 bp in length
* 57608 57707: gap of 100 bp
* 57708 58537: contig of 830 bp in length
* 58538 58637: gap of 100 bp
* 58638 59456: contig of 819 bp in length
* 59457 59556: gap of 100 bp
* 59557 60401: contig of 845 bp in length
* 60402 60501: gap of 100 bp
* 60502 61345: contig of 844 bp in length
* 61346 61445: gap of 100 bp
* 61446 62264: contig of 819 bp in length
* 62265 62364: gap of 100 bp
* 62365 63194: contig of 830 bp in length
* 63195 63294: gap of 100 bp
* 63295 64128: contig of 834 bp in length
* 64129 64228: gap of 100 bp

```

```

Query Match      68.7%; Score 20.6; DB 2; Length 78036;
Best Local Similarity 85.2%; Pred. No. 43;
Matches 23; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 2 ACAATGTATGTCGGGTACATCTATG 28
Db 40150 AAAACGTATGTCAGTGTACATTATG 40176

```

```

Search completed: December 6, 2004, 22:47:35
Job time : 142.012 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 15:16:44 ; Search time 16.5905 Seconds
(without alignments)
9492.325 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgatgtccgggtacatctatgac 30

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|---------------------|
| 1 | 30 | 100.0 | 30 | AAA37981 | Aaa37981 PCR prime |
| 2 | 30 | 100.0 | 2048 | AAA37962 | Aaa37962 Plasmid # |
| 3 | 30 | 100.0 | 2048 | AAA37963 | Aaa37963 Plasmid # |
| 4 | 30 | 100.0 | 2052 | AAA37959 | Aaa37959 Carrot pr |
| 5 | 30 | 100.0 | 2052 | AAA37961 | Aaa37961 Carrot pr |
| 6 | 30 | 100.0 | 2056 | AAA37964 | Aaa37964 Plasmid # |
| 7 | 23.6 | 78.7 | 2042 | AAV15144 | AAV15144 New promo |
| 8 | 19.6 | 65.3 | 2931 | ADN73150 | Adn73150 Thale cre |
| 9 | 19 | 63.3 | 2700 | ADB53615 | ADB53615 Primary r |
| 10 | 19 | 63.3 | 35042 | AAA81454 | AAA81454 N mening |
| 11 | 19 | 63.3 | 110000 | AAA81489_1 | Continuation (2 of |
| 12 | 19 | 63.3 | 236303 | AAAS11614 | AAAS11614 Human gen |
| 13 | 19 | 63.3 | 349980 | AAAF21611 | AAAF21611 Neisseria |
| 14 | 18.8 | 62.7 | 1630 | AAAG2657 | AAAG2657 cDNA sequ |
| 15 | 18.6 | 62.0 | 1569 | AAA30829 | AAA30829 Zebrafish |
| 16 | 18.6 | 62.0 | 1671 | AAA48445 | AAA48445 Zebrafish |
| 17 | 18.6 | 62.0 | 1671 | ADC42329 | ADC42329 cDNA enco |
| 18 | 18.6 | 62.0 | 2152 | AAA48446 | AAA48446 Zebrafish |
| 19 | 18.6 | 62.0 | 2152 | ADC42306 | ADC42306 cDNA enco |
| 20 | 18.6 | 62.0 | 2152 | ADH61060 | ADH61060 Zebrafish |
| 21 | 18.6 | 62.0 | 90442 | ADA03077 | Ada03077 Mouse mCG |

| | | | | | | |
|----|------|------|--------|----|-------------|--------------------|
| 22 | 18.6 | 62.0 | 90442 | 9 | ADA66361 | Ada66361 Mouse mCG |
| 23 | 18.6 | 62.0 | 90442 | 10 | ADB72815 | Adb72815 Mouse mCG |
| 24 | 18.6 | 62.0 | 90442 | 10 | ADC26997 | Adc26997 Mouse car |
| 25 | 18.6 | 62.0 | 90442 | 11 | ADL27155 | Adl27155 Mouse gen |
| 26 | 18.4 | 61.3 | 3177 | 4 | ABL28637 | AbL28637 Drosophil |
| 27 | 18.4 | 61.3 | 4354 | 4 | ABL11038 | AbL11038 Drosophil |
| 28 | 18.4 | 61.3 | 5460 | 4 | ABL28636 | AbL28636 Drosophil |
| 29 | 18.2 | 60.7 | 4919 | 4 | ABL15204 | AbL15204 Drosophil |
| 30 | 18.2 | 60.7 | 40050 | 9 | ADA02585 | Ada02585 Mouse Ics |
| 31 | 18.2 | 60.7 | 40050 | 10 | ADB72323 | Adb72323 Mouse Ics |
| 32 | 18.2 | 60.7 | 40050 | 10 | ADE95833 | Ade95833 Mouse Ics |
| 33 | 18 | 60.0 | 287 | 12 | ADL11418 | AdL11418 Cat flea |
| 34 | 18 | 60.0 | 287 | 12 | ADL11392 | AdL11392 Cat flea |
| 35 | 18 | 60.0 | 386 | 12 | ADL11486 | AdL11486 Cat flea |
| 36 | 18 | 60.0 | 401 | 4 | AAK96285 | Aak96285 Human neu |
| 37 | 18 | 60.0 | 401 | 4 | AAK97778 | Aak97778 Human neu |
| 38 | 18 | 60.0 | 401 | 6 | ABT01055 | ABT01055 Human neu |
| 39 | 18 | 60.0 | 401 | 6 | ABT02548 | ABT02548 Human neu |
| 40 | 18 | 60.0 | 448 | 12 | ADL10839 | AdL10839 Cat flea |
| 41 | 18 | 60.0 | 83709 | 12 | ADQ19964 | Adq19964 Human sof |
| 42 | 18 | 60.0 | 110000 | 4 | AAK95240_10 | Continuation (11 o |
| 43 | 18 | 60.0 | 110000 | 4 | AAK95240_11 | Continuation (12 o |
| 44 | 18 | 60.0 | 110000 | 4 | AAK96733_10 | Continuation (11 o |
| 45 | 18 | 60.0 | 110000 | 4 | AAK96733_11 | Continuation (12 o |

ALIGNMENTS

RESULT 1

AAA37981

ID AAA37981 standard; DNA; 30 BP.

XX

AC AAA37981;

XX

DT 18-AUG-2000 (first entry)

XX

DE PCR primer S used for carrot promoter amplification.

XX

KW Carrot; promoter; terminator; transgenic plant; breeding; fertility;

XX

XX PCR primer; ss.

OS Daucus carota.

XX

PN WO200020613-A1.

XX

PD 13-APR-2000.

XX

PF 28-SEP-1999; 99WO-JP005303.

XX

PR 02-OCT-1998; 98JP-00281124.

XX

PA (SUMO) SUMITOMO CHEM CO LTD.

PI

Nishikawa S, Oeda K;

DR

WPI; 2000-303791/26.

PT

New Plant promoters and terminators from Daucus carota L., useful in plant breeding, for e.g. controlling fertilities of plants.

XX

PS Example 6; Page 44; 81pp; English.

XX

This sequence represents a PCR primer used to amplify a carrot promoter sequence. The invention relates to plant promoters and terminators from Daucus carota L. which are capable of expressing a gene of interest in plants. The invention also includes a chimeric gene characterized in that it comprises the promoter and a desired gene linked to each other in the form capable of functioning. A method of producing a transformant comprising introducing the promoter, the chimeric gene or a vector comprising the promoter and a desired gene or terminator sequence into a host cell. The plant promoters and terminators are useful in plant breeding, for e.g. fertilities of plants may be controlled by expressing,

```

CC in the host cells, a sense or antisense gene of a male sterility related
CC Gene such as S-locus-specific RNase gene
XX
SQ Sequence 30 BP; 9 A; 6 C; 6 G; 9 T; 0 U; 0 Other;

Query Match      100.0%; Score 30; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.00016;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGCGGTGTACATCTATGAC 30
   |||||
Db 1 AACAAATGATGTCGCGGTGTACATCTATGAC 30

RESULT 2
AAA37962
ID AAA37962 standard; DNA; 2048 BP.
XX
AC AAA37962;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #1 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 71-73; 81pp; English.
XX
This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match      100.0%; Score 30; DB 3; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGCGGTGTACATCTATGAC 30
   |||||
Db 1742 AACAAATGATGTCGCGGTGTACATCTATGAC 1771

RESULT 4
AAA37959
ID AAA37959 standard; DNA; 2052 BP.
XX
AC AAA37959;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #1.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX

```

```

ID AAA37963 standard; DNA; 2048 BP.
XX
AC AAA37963;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #2 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
DR WPI; 2000-303791/26.
XX
New Plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 73-74; 81pp; English.
XX
This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2048 BP; 735 A; 317 C; 316 G; 680 T; 0 U; 0 Other;

Query Match      100.0%; Score 30; DB 3; Length 2048;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACAAATGATGTCGCGGTGTACATCTATGAC 30
   |||||
Db 1742 AACAAATGATGTCGCGGTGTACATCTATGAC 1771

RESULT 4
AAA37959
ID AAA37959 standard; DNA; 2052 BP.
XX
AC AAA37959;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #1.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX

```


PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
PI Nishikawa S, Oeda K;
XX WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 69-70; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 737 A; 317 C; 316 G; 682 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 30; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AACAAATGATGTCGGGTGACATCTATGAC 30
DB 1746 AACAAATGATGTCGGGTGACATCTATGAC 1775
XX
RESULT 5
AAA37961
ID AAA37961 standard; DNA; 2052 BP.
XX
AC AAA37961;
XX
DT 18-AUG-2000 (first entry)
XX
DE Carrot promoter sequence #2.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Claim 1; Page 78-79; 81pp; English.
XX
CC This sequence represents a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable

CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2052 BP; 736 A; 318 C; 317 G; 681 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 30; DB 3; Length 2052;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 AACAAATGATGTCGGGTGACATCTATGAC 30
DB 1746 AACAAATGATGTCGGGTGACATCTATGAC 1775
XX
RESULT 6
AAA37964
ID AAA37964 standard; DNA; 2056 BP.
XX
AC AAA37964;
XX
DT 18-AUG-2000 (first entry)
XX
DE Plasmid #3 DNA sequence used in mutation of promoter sequence.
XX
KW Carrot; promoter; terminator; transgenic plant; breeding; fertility; ds.
XX
OS Daucus carota.
XX
PN WO200020613-A1.
XX
PD 13-APR-2000.
XX
PF 28-SEP-1999; 99WO-JP005303.
XX
PR 02-OCT-1998; 98JP-00281124.
XX
PA (SUMO) SUMITOMO CHEM CO LTD.
XX
PI Nishikawa S, Oeda K;
XX
WPI; 2000-303791/26.
XX
PT New plant promoters and terminators from Daucus carota L., useful in
PT plant breeding, for e.g. controlling fertilities of plants.
XX
PS Example 8; Page 74-76; 81pp; English.
XX
CC This sequence represents a plasmid sequence used in a method for
CC introducing a mutation into a carrot promoter. The invention relates to
CC plant promoters and terminators from Daucus carota L. which are capable
CC of expressing a gene of interest in plants. The invention also includes a
CC chimeric gene characterized in that it comprises the promoter and a
CC desired gene linked to each other in the form capable of functioning. A
CC method of producing a transformant comprises introducing the promoter,
CC the chimeric gene or a vector comprising the promoter and a desired gene
CC or terminator sequence into a host cell. The plant promoters and
CC terminators are useful in plant breeding, for e.g. fertilities of plants
CC may be controlled by expressing, in the host cells, a sense or antisense
CC gene of a male sterility related gene such as S-locus-specific RNase gene
XX
SQ Sequence 2056 BP; 737 A; 319 C; 318 G; 682 T; 0 U; 0 Other;
XX
Query Match 100.0%; Score 30; DB 3; Length 2056;
Best Local Similarity 100.0%; Pred. No. 0.00036;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 AACAAATGATGTCGGGTGACATCTATGAC 30
Db 1746 AACAAATGATGTCGGGTGACATCTATGAC 1775

RESULT 7
AAV15144
ID AAV15144 standard; DNA; 2042 BP.
XX AC AAV15144;
XX DT 02-JUL-1998 (first entry)
XX DE New promoter used for root-specific expression in plants.
XX KW Promoter; root; carrot; Kuroda Gosun; root-specific expression;
KW plant cell; soil pathogen; improve; nutritive value; edible root plant;
XX ss.
XX OS Daucus carota.
XX PN EP824150-A2.
XX PD 18-FEB-1998.
XX PF 12-AUG-1997; 97EP-00113923.
XX PR 12-AUG-1996; 96JP-00212680.
XX PA (SUMO) SUMITOMO CHEM CO LTD.
XX PI Torikai S, Oeda K;
XX DR WPI; 1998-122310/12.
XX PT New carrot root gene, promoter and terminator - useful in genetic
PT engineering for directing root-specific gene expression.
XX PS Claim 2; Page 15-16; 31pp; English.
XX CC The present sequence represents a novel promoter, and is isolated from
CC the roots of carrots, cultivar Kuroda Gosun. The promoter can be used to
CC direct root-specific expression in plant cells. Since the promoter
CC enables expression of a desired protein in the roots of a plant, it is
CC useful in combat against pathogenic soil fungi and pests which are
CC difficult to kill by chemicals. It can also be used to improve the
CC nutritive value of edible root plants
XX SQ Sequence 2042 BP; 719 A; 356 C; 323 G; 644 T; 0 U; 0 Other;

Query Match 78.7%; Score 23.6; DB 2; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.44;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 AACAAATGATGTCGGGTGACATCTATGAC 30
Db 1737 AACAAATGATGTCGGGTGATATTATGAC 1766

RESULT 8
ADN73150/c
ID ADN73150 standard; cDNA; 2931 BP.
XX AC ADN73150;
XX DT 15-JUL-2004 (first entry)
XX DE Thale cress cDNA upregulated in E2Fa/Dpa expressing plants SeqID 1045.
XX KW gene; ss; plant; transgenic; E2Fa/Dpa transcription factor;
KW growth regulator; animal feed product; thale cress;
KW cell wall biosynthesis; nitrogen metabolism; carbon metabolism.
XX PN WO2003065993-A2.

QY 5 ATGTATGTCGGGTGACATCTATGAC 30
Db 2591 AAGTACGTCGTCGTACATCTATGAC 2566

Query Match 65.3%; Score 19.6; DB 12; Length 2931;
Best Local Similarity 84.6%; Pred. No. 40;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

RESULT 9
ADB53615/c
ID ADB53615 standard; DNA; 2700 BP.
XX AC ADB53615;
XX DT 04-DEC-2003 (first entry)
XX DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4157.
XX KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX OS Rattus norvegicus.
XX PN WO2003065993-A2.

QY 5 ATGTATGTCGGGTGACATCTATGAC 30
Db 2591 AAGTACGTCGTCGTACATCTATGAC 2566

Query Match 65.3%; Score 19.6; DB 12; Length 2931;
Best Local Similarity 84.6%; Pred. No. 40;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Arabidopsis thaliana.
WO2004035798-A2.
29-APR-2004.
20-OCT-2003; 2003WO-EP011658.
18-OCT-2002; 2002EP-00079408.
(CROP-) CROPDESIGN NV.
Inze D, De Veylder L, Vlieghe K;
WPI; 2004-348466/32.
P-PSDB; ADN73151.
Altering plant characteristics, useful for producing plants for enzyme or
pharmaceutical production comprises modifying in a plant, expression of
one or more nucleic acids and/or modifying level or activity of one or
more proteins.
Claim 1; SEQ ID NO 1045; 134pp; English.
This invention relates to a novel method for altering one or more plant
characteristics. Specifically, it refers to identifying genes that are up
- or down-regulated in transgenic plants overexpressing the heterodimeric
E2Fa/Dpa transcription factor of Arabidopsis and using these sequences to
alter plant characteristics accordingly. The present invention describes
generating transgenic plants for the production of growth regulators,
enzymes, therapeutics, pharmaceuticals and animal feed products, where
the altered plant characteristics are selected from increased yield or
biomass, enhanced survival capacity, stress tolerance, plant architecture
or physiology, altered endoreplication, biochemistry, signal
transduction, storage lipid mobilisation and/or altered photosynthesis,
each relative to the corresponding wild type plants. Accordingly, these
sequences can also be useful as positive or negative selectable markers
during transformation of cells or tissues. The identified genes play a
role in a variety of biological processes such as DNA replication, cell
wall biosynthesis, nitrogen and/or carbon metabolism or they function as
transcription factors. This polynucleotide sequence is thale cress cDNA
upregulated 1.3 fold or more in plants overexpressing the E2Fa/Dpa
transcription factor, given in an exemplification of the invention.
SQ Sequence 2931 BP; 818 A; 547 C; 682 G; 884 T; 0 U; 0 Other;

```


WP AAA81489_2 200001 310000
 WP AAA81489_3 300001 410000
 WP AAA81489_4 400001 510000
 WP AAA81489_5 500001 610000
 WP AAA81489_6 600001 710000
 WP AAA81489_7 700001 810000
 WP AAA81489_8 800001 837096

Query Match 63.3%; Score 19; DB 3; Length 110000;

Best Local Similarity 81.5%; Pred. No. 1.6e+02;

Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AACRATGATGTCGGTGTCATCTAT 27

Db 16813 AACAAATGATTTCTATTGTTATCTAT 16839

RESULT 12

AAS11614

ID AAS11614 standard; DNA; 236303 BP.

XX AC AAS11614;

XX DT 24-OCT-2001 (first entry)

DE Human genomic DNA containing exons 2-17 of the CRIM1 gene.

XX KW CRIM-1; Human; human chromosome 2p21-16.3; ophthalmological;
 KW neuroprotective; renal; osteopathic; dental; vulnarary; immunogen;
 KW antibody; gene therapy; neurodegenerative disease; eye disorder;
 KW cataract; bone morphogenic protein; BMG; renal disease; bone abnormality;
 KW tooth abnormality; wound; ds.

XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT exon 33104..33277

FT /tag= a

FT /number= 2

FT 33278..77746

FT /tag= b

FT /number= 2

FT 77747..77989

FT /tag= c

FT /number= 3

FT 77990..79103

FT /tag= d

FT /number= 3

FT 79104..79224

FT /tag= e

FT /number= 4

FT 79225..101022

FT /tag= f

FT /number= 4

FT 101023..101144

FT /tag= g

FT /number= 5

FT 101145..113377

FT /tag= h

FT /number= 5

FT 113378..113560

FT /tag= i

FT /number= 6

FT 113561..115985

FT /tag= j

FT /number= 6

FT 115986..116183

FT /tag= k

FT /number= 7

FT 116184..135707

FT /tag= l

FT /number= 7

FT 135708..135836

FT intron /tag= m
 FT /number= 8
 FT 135837..146471
 FT /tag= m
 FT /number= 8
 FT 146472..146628
 FT /tag= o
 FT /number= 9
 FT 146629..148761
 FT /tag= p
 FT /number= 9
 FT 148762..148883
 FT /tag= q
 FT /number= 10
 FT 148884..150044
 FT /tag= r
 FT /number= 10
 FT 150045..150254
 FT /tag= s
 FT /number= 11
 FT 150255..153815
 FT /tag= t
 FT /number= 11
 FT 153816..154031
 FT /tag= u
 FT /number= 12
 FT 154032..158580
 FT /tag= v
 FT /number= 12
 FT 158581..158802
 FT /tag= w
 FT /number= 13
 FT 158803..173982
 FT /tag= x
 FT /number= 13
 FT 173983..174177
 FT /tag= y
 FT /number= 14
 FT 174178..181006
 FT /tag= z
 FT /number= 14
 FT 181007..181129
 FT /tag= aa
 FT /number= 15
 FT 181130..183612
 FT /tag= ab
 FT /number= 15
 FT 183613..183800
 FT /tag= ac
 FT /number= 16
 FT 183801..185152
 FT /tag= ad
 FT /number= 16
 FT 185153..187765
 FT /tag= ae
 FT /number= 17
 FT WO200138519-A1.
 PN 31-MAY-2001.
 PD 24-NOV-2000; 2000WO-AU001435.
 PF 26-NOV-1999; 99AU-00004348.
 PR (UYQU) UNIV QUEBENS LAND.
 PA Little M, Yamada T, Holmes G, Georgas K, Kolle G, Wilkinson L;
 PI WPI; 2001-343951/36.
 DR Nucleic acids from human chromosome 2p21-16.3 and the encoded peptide,
 PT useful for preventing, diagnosing and treating e.g. eye disease,

PT especially cataract formation.
 XX Claim 4; Fig 3; 169pp; English.
 XX
 CC The invention relates to nucleic acids from human chromosome 2p21-16.3
 CC and the encoded peptide (and mouse and chicken orthologues) that
 CC comprises a PGECCPLP group, an insulin-like growth factor binding protein
 CC (IGFBP)-like domain, cysteine-rich domains, an RGD (undefined) group and
 CC a transmembrane domain. The protein, e.g. CRIM1, interacts with peptides
 CC of the transforming growth factor superfamily. A composition comprising
 CC an expression construct comprising the nucleic acids of the invention or
 CC a mimetic which antagonises or mimics an activity of a CRIM1 polypeptide
 CC may be used in a method for modulating the biological activity of a
 CC polypeptide of the bone morphogenic protein (BMP) family. In this way
 CC they may be used to prevent or treat an eye disease, especially cataract
 CC formation. They may also be used to treat neurodegenerative diseases,
 CC renal and kidney disease, bone and tooth abnormalities, wounds and skin
 CC damage, e.g. by use of the nucleic acid in gene therapy by using
 CC antibodies directed against CRIM1 polypeptides. The present sequence is a
 CC Human genomic DNA containing exons 2-17 of the CRIM1 gene
 XX
 SQ Sequence 236303 BP; 70199 A; 46544 C; 47996 G; 71563 T; 0 U; 1 Other;
 Query Match 63.3%; Score 19; DB 4; Length 236303;
 Best Local Similarity 81.5%; Pred. No. 1.8e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 4 AATGTATGTCGGTGATACATCTATGAC 30
 DB 133196 AATGTGTGATGATGATACATCTGTGAC 133222
 RESULT 13
 AAF21611
 ID AAF21611 standard; DNA; 349980 BP.
 XX
 AC AAF21611;
 XX
 DT 13-MAR-2001 (first entry)
 XX
 DE Neisseria meningitidis B nucleotide sequence SEQ ID NO:112.
 XX
 KW Neisseria meningitidis; Neisseria gonorrhoeae; immunogenic; vaccine;
 KW diagnosis; antigen; detection; infection; gene therapy; antibacterial;
 KW ds.
 XX
 OS Neisseria meningitidis.
 XX
 PN WO200066791-A1.
 XX
 PD 09-NOV-2000.
 XX
 PF 08-MAR-2000; 2000WO-US005928.
 XX
 PR 30-APR-1999; 99US-0132068P.
 PR 08-OCT-1999; 99WO-US023573.
 PR 28-FEB-2000; 2000GB-00004695.
 XX
 PA (CHIR) CHIRON CORP.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Pizza M, Hickey B, Peterson J, Tettelin H, Venter JC;
 PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
 PI Rappuoli R, Frazer CM, Grandi G;
 XX
 DR WPI; 2000-647603/62.
 XX
 PT Neisseria meningitidis B full length genome sequence and open reading
 PT frames are used to detect, treat and prevent Neisserial infections.
 XX
 PS Claim 7; Appendix A; 692pp; English.
 XX
 CC The present invention describes the full length genome of Neisseria

CC meningitidis B (NMB). The sequences in AAF21544 and AAF21607 to AAF21613
 CC represent fragments of the NMB genomic sequence, as the sequence was too
 CC long to go in a record on its own it was split into 8 sequences which
 CC overlap each other at the beginning and end of each sequence by 49980 bp
 CC (i.e. the last 49980 bp of AAF21544 is repeated at the beginning of
 CC AAF21607, the last 49980 bp of AAF21607 are repeated at the beginning of
 CC AAF21608, and so on). AAF21545 to AAF21588 encode the Neisseria proteins
 CC given in AAB58550 to AAB58593, and AAF21589 to AAF21606 represent PCR
 CC primers which are used in the exemplification of the present invention.
 CC The NMB genome and fragments from it have antibacterial activity, and can
 CC be used in vaccines and gene therapy. Neisseria nucleic acids, proteins
 CC and/or antibodies which binds to the proteins can be used in compositions
 CC for treating or preventing infection due to Neisserial bacteria or as a
 CC diagnostic reagent for detecting the presence of Neisserial bacteria or
 CC of antibodies raised to Neisserial bacteria. Computers, computer memory,
 CC computer storage medium or computer databases can be used in a search to
 CC identify open reading frames (ORFs) or coding sequences within the NMB
 CC genome. The DNA sequences provide further opportunities to find antigenic
 CC or immunogenic proteins which are more effective in vaccines than the
 CC outer membrane proteins currently used
 XX
 SQ Sequence 349980 BP; 87189 A; 93501 C; 84627 G; 84663 T; 0 U; 0 Other;
 Query Match 63.3%; Score 19; DB 3; Length 349980;
 Best Local Similarity 81.5%; Pred. No. 2e+02;
 Matches 22; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1 AACATGTATGTCGGTGATACATCTAT 27
 DB 52042 AACATGTATTTCTATTGTCATCTAT 52068
 RESULT 14
 AAS62657/C
 ID AAS62657 standard; cDNA; 1630 BP.
 XX
 AC AAS62657;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE cDNA sequence #444 encoding novel human secreted protein.
 XX
 KW Human secreted protein; hyperproliferative disorder; autoimmune disorder;
 KW immune deficiency disorder; blood disorder; inflammatory disorder;
 KW infectious disorder; gene therapy; antimicrobial; hepatotropic;
 KW immunosuppressive; antirheumatic; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200177291-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 29-MAR-2001; 2001WO-US010485.
 XX
 PR 06-APR-2000; 2000US-0195604P.
 XX
 PA (GENY) GENETICS INST INC.
 XX
 PI Wong GG, Clark HF, Fechtel K, Agoatino MJ, Howes SH, Resnick RJ;
 PI Gulukota K, Graham JR;
 XX
 DR WPI; 2002-010900/01.
 XX
 PT New polynucleotides encoding secreted proteins useful for treating e.g.
 PT asthma, HIV and Crohn's disease.
 XX
 PS Claim 1; Page 308; 391pp; English.
 XX
 CC The present invention relates to the isolation of novel cDNA sequences
 CC which encode human secreted proteins. The cDNA sequences have been
 CC derived from a variety of human tissues. The invention also provides a
 CC method for producing proteins from these polynucleotide sequences. The

CC proteins are useful for identifying compounds that modulate their
 CC activity and production, and the cell is also useful for identifying
 CC compounds that modulate expression of the polynucleotide sequences
 CC encoding the secreted proteins. The sequences of the invention are useful
 CC for treating diseases such as hyperproliferative disorders (e.g. cancer),
 CC immune deficiency disorders (e.g. severe combined immunodeficiency
 CC (SCID)), autoimmune disorders (e.g. multiple sclerosis), blood disorders
 CC (e.g. thrombocytopenia), inflammatory disorders (e.g. arthritis) and
 CC infectious disorders (e.g. hepatitis). The polynucleotide sequences of
 CC the invention are also useful in gene therapy. AAS62214-AAS62838
 CC represent the cDNA sequences of the invention that encode for novel human
 CC secreted proteins

XX
 SQ Sequence 1630 BP; 409 A; 346 C; 284 G; 590 T; 0 U; 1 Other;
 Query Match 62.7%; Score 18.8; DB 6; Length 1630;
 Best Local Similarity 76.7%; Pred. No. 87;
 Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 AACATGTATGTCGGGTGATCATCTATGAC 30
 |||||
 Db 633 AACGAAGTAAGCCCTGTGATCATCTATGAC 604
 |||||

RESULT 15

AAA30829
 ID AAA30829 standard; DNA; 1569 BP.

XX AC AAA30829;

DT 15-SEP-2003 (revised)

XX 29-AUG-2000 (first entry)

DE Zebrafish PTH3R receptor coding sequence.

XX Zebrafish; PTH1R receptor; PTH3R receptor; diagnosis; cancer;
 KW parathyroid hormone type 1 receptor; parathyroid hormone type 3 receptor;
 KW ds.

XX OS Danio rerio.

XX Key Location/Qualifiers
 FH CDS 1..1569
 FT /tag= a
 FT /product= "PTH3R receptor"
 FT /partial
 FT /note= "no stop codon given"

XX PN WO200032771-A1.

XX PD 08-JUN-2000.

XX PF 28-MAY-1999; 99WO-US011883.

XX PR 30-NOV-1998; 98US-0110467P.

XX PA (JUEP/) JUEPPNER H.

XX PA (RUBI/) RUBIN D A.

XX PI Jueppner H, Rubin DA;

XX WPI, 2000-412319/35.

XX P-FSDB; AAY90231.

XX Novel zebrafish parathyroid hormone/parathyroid hormone related peptide
 PT receptor 3 and isolated nucleic acid encoding zebrafish parathyroid
 PT hormone receptor 1 for treating disorders associated with receptor
 PT function.

XX Claim 23; Fig 1d; 11pp; English.

XX This sequence encodes a parathyroid hormone receptor type 3 (PTH3R)
 CC receptor protein of the invention. The invention also relates to a PTH1R

CC receptor protein. Antagonists of PTH1R or PTH3R can be used for the
 CC treatment of diseases associated with an increase in PTH1R or PTH3R
 CC activity, respectively. The peptides are used for diagnosis or prognosis
 CC of diseases and disorders associated with PTH3R or PTH1R, such as cancer.
 CC The polypeptides can be used as a molecular weight markers on sodium
 CC dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) gels, or
 CC on molecular sieve gel filtration columns. Antigenic epitope-bearing
 CC peptides and polypeptides are useful to raise antibodies, including
 CC monoclonal antibodies, that bind specifically to a polypeptide. The
 CC peptides are useful during diagnosis of diseases and disorders in mammals
 CC involving PTH1R or PTH3R receptor expression or function. Mutations that
 CC affect PTH1R or PTH3R sequence and/or expression levels of PTH1R or PTH3R
 CC could be diagnostic for patients with disease or disorders of a
 CC developmental, physiological or neurological nature. The nucleic acid
 CC molecules are valuable for chromosome identification. The mapping of DNAs
 CC to chromosomes is an important first step in correlating those sequences
 CC with genes associated with disease. (Updated on 15-SEP-2003 to
 CC standardise OS field)

XX SQ Sequence 1569 BP; 382 A; 358 C; 418 G; 411 T; 0 U; 0 Other;
 Query Match 62.0%; Score 18.6; DB 3; Length 1569;
 Best Local Similarity 84.0%; Pred. No. 1.1e+02;
 Matches 21; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 6 TGTATGTCGGGTGATCATCTATGAC 30
 |||||

Db 246 TCTGTGTCCTGAGTACATCTATGAC 270
 |||||

Search completed: December 6, 2004, 19:02:31
 Job time : 19.5905 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 18:08:52 ; Search time 3.00716 Seconds
(without alignments)
7090.965 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 acaatgtatgtccggtgtacatcatgac 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS-COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 23.6 | 78.7 | 2042 | 2 | US-08-911-434A-2 |
| 2 | 18.6 | 62.0 | 2152 | 4 | US-09-449-632-3 |
| 3 | 17.8 | 59.3 | 438 | 3 | US-08-651-155B-134 |
| 4 | 17.8 | 59.3 | 438 | 4 | US-09-194-036B-134 |
| 5 | 17.8 | 59.3 | 908 | 3 | US-09-457-046B-15 |
| 6 | 17.4 | 58.0 | 198 | 4 | US-09-248-796A-7260 |
| 7 | 17.4 | 58.0 | 1507 | 3 | US-08-605-150A-15 |
| 8 | 17.4 | 58.0 | 2182 | 4 | US-09-221-017B-916 |
| 9 | 17.4 | 58.0 | 2662 | 3 | US-08-750-357-8 |
| 10 | 17.4 | 58.0 | 3848 | 3 | US-09-112-096-28 |
| 11 | 17.4 | 58.0 | 5668 | 3 | US-09-112-096-14 |
| 12 | 17.4 | 58.0 | 5668 | 4 | US-09-636-215-777 |
| 13 | 17.4 | 58.0 | 5668 | 4 | US-09-685-166A-777 |
| 14 | 17.4 | 58.0 | 5668 | 4 | US-09-679-426-777 |
| 15 | 17.4 | 58.0 | 48974 | 3 | US-08-920-422-17 |
| 16 | 17.2 | 57.3 | 351 | 4 | US-09-328-352-182 |
| 17 | 17.2 | 57.3 | 5474 | 1 | US-08-137-252-2 |
| 18 | 17.2 | 57.3 | 11663 | 1 | US-08-446-932-1 |
| 19 | 17.2 | 57.3 | 11663 | 1 | US-08-801-263A-1 |
| 20 | 17.2 | 57.3 | 11663 | 1 | US-08-801-263A-7 |
| 21 | 17.2 | 57.3 | 11663 | 3 | US-09-102-248-1 |
| 22 | 17.2 | 57.3 | 11663 | 3 | US-09-102-248-7 |
| 23 | 17.2 | 57.3 | 11663 | 4 | US-09-167-764-1 |
| 24 | 17.2 | 57.3 | 11663 | 4 | US-09-367-764-7 |
| 25 | 17 | 56.7 | 745 | 4 | US-09-270-767-12913 |
| 26 | 17 | 56.7 | 11580 | 3 | US-09-334-220-4 |
| 27 | 16.8 | 56.0 | 275 | 1 | US-07-789-919A-1 |

28 16.8 56.0 275 1 US-08-209-846A-1 Sequence 1, Appli
29 16.8 56.0 275 2 US-08-472-809B-1 Sequence 1, Appli
30 16.8 56.0 275 3 US-08-438-265-1 Sequence 1, Appli
31 16.8 56.0 1529 3 US-08-821-278A-1 Sequence 1, Appli
32 16.8 56.0 1609 4 US-09-449-632-1 Sequence 1, Appli
33 16.8 56.0 2166 4 US-09-107-532A-3235 Sequence 3235, Ap
34 16.8 56.0 4819 2 US-09-774-528-72 Sequence 72, Appli
35 16.8 56.0 5707 2 US-08-472-809B-8 Sequence 8, Appli
36 16.8 56.0 6345 2 US-08-472-809B-7 Sequence 7, Appli
37 16.8 56.0 640681 4 US-09-790-988-1 Sequence 1, Appli
38 16.4 54.7 148 4 US-09-513-999C-30057 Sequence 30057, A
39 16.4 54.7 484 3 US-08-945-983-3 Sequence 3, Appli
40 16.4 54.7 484 3 US-08-945-983-5 Sequence 5, Appli
41 16.4 54.7 600 6 5198345-4 Patent No. 5198345
42 16.4 54.7 610 4 US-09-270-767-6688 Sequence 6688, Ap
43 16.4 54.7 610 4 US-09-270-767-21970 Sequence 21970, A
44 16.4 54.7 3349 4 US-09-375-318-36 Sequence 36, Appli
45 16.4 54.7 51259 3 US-08-781-891-209 Sequence 209, App

ALIGNMENTS

RESULT 1
US-08-911-434A-2
; Sequence 2, Application US/08911434A
; Patent No. 5959176
; GENERAL INFORMATION:
; APPLICANT: TORIKAI, Satomi
; APPLICANT: OEDA, Kenji
; TITLE OF INVENTION: PLANT PROMOTER AND UTILIZATION THEREOF
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
; STREET: P.O. BOX 747
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 22040

COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,434A
FILING DATE: 12-AUG-1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Raymond C.
REGISTRATION NUMBER: 21,066
REFERENCE/DOCKET NUMBER: 2185-0199P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)205-8000
TELEFAX: (703)205-8050
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Daucus carota L.
INDIVIDUAL ISOLATE: Kuroda Gosun
FEATURE:
NAME/KEY: promoter
LOCATION: 1..2042

Query Match 78.7%; Score 23.6; DB 2; Length 2042;
Best Local Similarity 86.7%; Pred. No. 0.12;


```

; ANTI-SENSE: YES
; ORIGINAL SOURCE: DNA (other)
; SEQUENCE DESCRIPTION: SEQ ID NO: 134:
US-09-194-036B-134
Query Match 59.3%; Score 17.8; DB 4; Length 438;
Best Local Similarity 75.9%; Pred. No. 43;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ACAATGATGTCGGGTGATCATCTATGAC 30
DB 149 ATAATGAATGGCCGGTGATCAACCCATGAC 177

RESULT 5
US-09-457-046B-15
; Sequence 15, Application US/09457046B
; Patent No. 6287835
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney et al.
; TITLE OF INVENTION: Transacylases of the Paclitaxel Biosynthetic Pathway
; FILE REFERENCE: 53679
; CURRENT APPLICATION NUMBER: US/09/457,046B
; CURRENT FILING DATE: 1999-12-07
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 908
; TYPE: DNA
; ORGANISM: Taxis cuspidata
US-09-457-046B-15
Query Match 59.3%; Score 17.8; DB 3; Length 908;
Best Local Similarity 75.9%; Pred. No. 48;
Matches 22; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 ACAATGATGTCGGGTGATCATCTATGAC 30
DB 775 ACAGAGAATTCGGGTGAAGATCTATGAC 803

RESULT 6
US-09-248-796A-7260/c
; Sequence 7260, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 7260
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-7260
Query Match 58.0%; Score 17.4; DB 4; Length 198;
Best Local Similarity 77.8%; Pred. No. 58;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 AACATGATGTCGGGTGATCATCTAT 27
DB 196 AACATGAATGTCACCTTGATCATGCAT 170

RESULT 7
US-08-605-150A-15/c
; Sequence 15, Application US/08605150A
; Patent No. 6103520
; GENERAL INFORMATION:
; APPLICANT: Topfer, Reinhard
; APPLICANT: Hausmann, Ludger
; APPLICANT: Schell, Josef
; TITLE OF INVENTION: GLYCEROL-3-PHOSPHATE DEHYDROGENASE
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klein & Szekeres
; STREET: 4199 Campus Drive, Suite 700
; CITY: Irvine
; STATE: CA
; COUNTRY: USA
; ZIP: 92715
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/605,150A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP94/02936
; FILING DATE: 02-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P4329827.3
; FILING DATE: 03-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Szekeres, Gabor L.
; REGISTRATION NUMBER: 28,675
; REFERENCE/DOCKET NUMBER: 542-04-PA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 714-854-5502
; TELEFAX: 714-854-4897
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1507 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Cuphea lanceolata
; IMMEDIATE SOURCE:
; LIBRARY: Genomic lambda FIX II
; CLONE: C1GPDH99
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1193..1375
US-08-605-150A-15
Query Match 58.0%; Score 17.4; DB 3; Length 1507;
Best Local Similarity 77.8%; Pred. No. 81;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 3 CAATGTATGTCGGGTGATCATCTATGA 29
DB 1018 CAATGTATGTCCTCGTGCAAACTATGA 992

RESULT 8
US-09-221-017B-916/c
; Sequence 916, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:

```

```
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FASTSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 916:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2182 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: UNKNOWN
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..2182
US-09-221-017B-916

Query Match 58.0%; Score 17.4; DB 4; Length 2182;
Best Local Similarity 77.8%; Pred. No. 86;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 3 CAATGTATGTCGGGTACATCTATGA 29
Db 1032 CAATGTGGTCTGGAGTATATCCATGA 1006

RESULT 9
US-08-750-357-8/c
; Sequence 8, Application US/08750357
; Patent No. 6008437
; GENERAL INFORMATION:
; APPLICANT: KREBBERS, Enno
; APPLICANT: WILLIAMS, Mark
; APPLICANT: LEEMANS, Jan
; TITLE OF INVENTION: USE OF ANTHOCYANIN GENES TO MAINTAIN
```

```
; TITLE OF INVENTION: MALE STERILE PLANTS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/750,357
; FILING DATE: 21-MAR-1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 018030-100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2662 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-750-357-8

Query Match 58.0%; Score 17.4; DB 3; Length 2662;
Best Local Similarity 77.8%; Pred. No. 89;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ACAATGTATGTCGGGTACATCTATG 28
Db 392 AAAATGAATGTCGGTGGACTTCTAGG 366

RESULT 10
US-09-112-096-28/c
; Sequence 28, Application US/09112096
; Patent No. 6194152
; GENERAL INFORMATION:
; APPLICANT: Reiner Laus
; APPLICANT: Michael H. Shapiro
; APPLICANT: Larisa Tsavaler
; TITLE OF INVENTION: Prostate Tumor Polynucleotide and
; TITLE OF INVENTION: Antigen Compositions
; FILE REFERENCE: 7636-0015.30
; CURRENT APPLICATION NUMBER: US/09/112,096
; CURRENT FILING DATE: 1998-07-09
; EARLIER APPLICATION NUMBER: 60/056,110
; EARLIER FILING DATE: 1997-08-20
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 3848
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-112-096-28

Query Match 58.0%; Score 17.4; DB 3; Length 3848;
Best Local Similarity 77.8%; Pred. No. 94;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 2 ACAATGTATGTCGGGTACATCTATG 28
Db 2323 ACAATTATGTTCTTCTACATCTCTG 2297
```


; FILE REFERENCE: 210121.427C20
; CURRENT APPLICATION NUMBER: US/09/679,426
; CURRENT FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 895
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 777
; LENGTH: 5668
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-679-426-777

Query Match 58.0%; Score 17.4; DB 4; Length 5668;
Best Local Similarity 77.8%; Pred. No. 1e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ACAATGTATGTCGGGTACATCTATG 28
DB 4143 ACAATTATGTTCTTTCTACATCTCTG 4117

RESULT 15
US-08-920-422-17
; Sequence 17, Application US/08920422A
; Patent No. 6255473
; GENERAL INFORMATION:
; APPLICANT: Vitek, Michael P.
; APPLICANT: Mitsuda, No. 6255473iaki
; APPLICANT: Roses, Allen D.
; TITLE OF INVENTION: Presenilin-1 Gene Promoter
; FILE REFERENCE: VITEKPRESENILIN
; CURRENT APPLICATION NUMBER: US/08/920,422A
; CURRENT FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 17
; LENGTH: 48974
; TYPE: DNA
; ORGANISM: Mus musculus
US-08-920-422-17

Query Match 58.0%; Score 17.4; DB 3; Length 48974;
Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 21; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ACAATGTATGTCGGGTACATCTATG 28
DB 31022 ATAGTGTATGTCATGGACATGTATG 31048

Search completed: December 6, 2004, 22:52:28
Job time : 4.00716 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2004, 19:02:39 ; Search time 17.7259 Seconds
(without alignments)
9300.077 Million cell updates/sec

Title: US-09-806-197-24

Perfect score: 30

Sequence: 1 aacaatgatgtccgggtacatcatgac 30

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3695051 seqs, 2747533894 residues

Total number of hits satisfying chosen parameters: 7390102

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|------------------------|-------------------|
| C 1 | 19.6 | 65.3 | 380 | 17 US-10-437-963-26247 | Sequence 26247, A |
| C 2 | 19.4 | 64.7 | 374849 | 13 US-10-087-192-1627 | Sequence 1627, Ap |
| C 3 | 19.2 | 64.0 | 42772 | 13 US-10-087-192-1903 | Sequence 1903, Ap |
| C 4 | 19.2 | 64.0 | 2731748 | 17 US-10-297-465A-1 | Sequence 1, Appli |
| C 5 | 19 | 63.3 | 512 | 17 US-10-021-323-6008 | Sequence 6008, Ap |
| C 6 | 19 | 63.3 | 50000 | 14 US-10-152-742A-22 | Sequence 22, Appl |
| C 7 | 19 | 63.3 | 1601042 | 13 US-10-027-632-59064 | Sequence 59064, A |
| C 8 | 19 | 63.3 | 1601042 | 15 US-10-027-632-59064 | Sequence 59064, A |
| C 9 | 18.8 | 62.7 | 1630 | 9 US-09-822-830A-444 | Sequence 444, App |
| C 10 | 18.6 | 62.0 | 665 | 13 US-10-027-632-8033 | Sequence 8033, Ap |
| C 11 | 18.6 | 62.0 | 665 | 15 US-10-027-632-8033 | Sequence 8033, Ap |
| C 12 | 18.6 | 62.0 | 2152 | 15 US-10-372-095-3 | Sequence 3, Appli |

| | | | | | |
|------|------|------|---------|-------------------------|---------------------|
| 13 | 18.6 | 62.0 | 90442 | 14 US-10-105-637-1 | Sequence 1, Appli |
| 14 | 18.6 | 62.0 | 90442 | 15 US-10-034-650-43 | Sequence 43, Appl |
| C 15 | 18.4 | 61.3 | 255 | 16 US-10-424-599-132779 | Sequence 132779, Ap |
| C 16 | 18.4 | 61.3 | 650 | 13 US-10-027-632-320106 | Sequence 320106, Ap |
| C 17 | 18.4 | 61.3 | 650 | 15 US-10-027-632-320106 | Sequence 320106, Ap |
| C 18 | 18.4 | 61.3 | 662 | 13 US-10-027-632-84202 | Sequence 84202, A |
| C 19 | 18.4 | 61.3 | 662 | 15 US-10-027-632-84202 | Sequence 84202, A |
| C 20 | 18.4 | 61.3 | 188017 | 13 US-10-087-192-1951 | Sequence 1951, Ap |
| C 21 | 18.2 | 60.7 | 40050 | 16 US-10-052-482-91 | Sequence 91, Appli |
| C 22 | 18.2 | 60.7 | 310122 | 18 US-10-417-375-1 | Sequence 1, Appli |
| C 23 | 18 | 60.0 | 287 | 16 US-10-621-901-1833 | Sequence 1833, Ap |
| C 24 | 18 | 60.0 | 287 | 16 US-10-621-901-1859 | Sequence 1859, Ap |
| C 25 | 18 | 60.0 | 386 | 16 US-10-621-901-1927 | Sequence 1927, Ap |
| C 26 | 18 | 60.0 | 401 | 9 US-09-795-668-1084 | Sequence 1084, Ap |
| C 27 | 18 | 60.0 | 401 | 9 US-09-795-668-1084 | Sequence 1084, Ap |
| C 28 | 18 | 60.0 | 401 | 9 US-09-946-807-1084 | Sequence 1084, Ap |
| C 29 | 18 | 60.0 | 448 | 16 US-10-621-901-1276 | Sequence 1276, Ap |
| C 30 | 18 | 60.0 | 620 | 18 US-10-425-115-32472 | Sequence 32472, A |
| C 31 | 18 | 60.0 | 649 | 16 US-10-424-599-103647 | Sequence 103647, A |
| C 32 | 18 | 60.0 | 2194 | 17 US-10-437-963-21442 | Sequence 21442, A |
| C 33 | 18 | 60.0 | 2880 | 16 US-10-424-599-74550 | Sequence 74550, A |
| C 34 | 18 | 60.0 | 42007 | 15 US-10-085-117-337 | Sequence 337, App |
| C 35 | 18 | 60.0 | 1503841 | 9 US-09-795-668-1 | Sequence 1, Appli |
| C 36 | 18 | 60.0 | 1503841 | 9 US-09-795-668-1 | Sequence 1, Appli |
| C 37 | 18 | 60.0 | 1503841 | 9 US-09-946-807-1 | Sequence 1, Appli |
| C 38 | 17.8 | 59.3 | 141 | 13 US-10-033-528-1890 | Sequence 1890, Ap |
| C 39 | 17.8 | 59.3 | 141 | 15 US-10-039-926-1890 | Sequence 1890, Ap |
| C 40 | 17.8 | 59.3 | 421 | 17 US-10-437-963-89493 | Sequence 89493, A |
| C 41 | 17.8 | 59.3 | 428 | 14 US-10-198-846-922 | Sequence 922, App |
| C 42 | 17.8 | 59.3 | 500 | 10 US-09-918-995-23954 | Sequence 23954, A |
| C 43 | 17.8 | 59.3 | 662 | 14 US-10-198-846-829 | Sequence 829, App |
| C 44 | 17.8 | 59.3 | 881 | 14 US-10-198-846-10611 | Sequence 10611, A |
| C 45 | 17.8 | 59.3 | 908 | 9 US-09-866-572A-15 | Sequence 15, Appl |

ALIGNMENTS

RESULT 1

US-10-437-963-26247/c

; Sequence 26247, Application US/10437963

; Publication NO. US20040123343A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Kovalic, David K.

; APPLICANT: Zhou, Yihua

; APPLICANT: Cao, Yongwei

; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.

; APPLICANT: Barbazuk, Brad

; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963

; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 26247

; LENGTH: 380

; TYPE: DNA

; ORGANISM: Oryza sativa

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(380)

; OTHER INFORMATION: unsure at all n locations

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT_MRT4530_31056C.1

US-10-437-963-26247

Query Match 65.3%; Score 19.6; DB 17; Length 380;

Best Local Similarity 78.6%; Pred. No. 48;

Matches 22; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

RESULT 9
US-09-822-830A-444/c
; Sequence 444, Application US/09822830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulukota, Kamalakara
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29

RESULT 15
US-10-424-599-132779/c
; Sequence 132779, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:

THIS PAGE BLANK (USPTO)